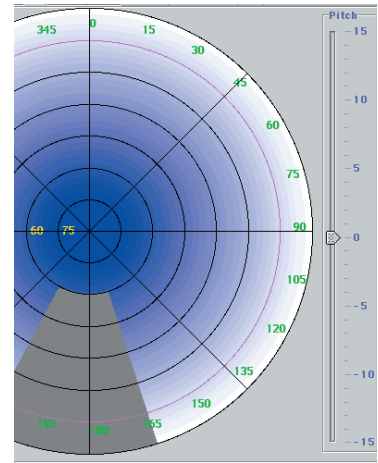


Communication Systems Technologies ■



Strategies for Optimizing Bandwidth Efficiency	211
Todd Landers (SSC San Diego)	
Tools for Analyzing and Describing the Impact of Superstructure Blockage on Availability in Shipboard and Submarine Satellite Communications Systems	219
Roy A. Axford, Jr. (SSC San Diego)	
Gerald B. Fitzgerald (The MITRE Corporation)	
Advanced Enclosed Mast/Sensor (AEM/S) System	230
John H. Meloling (SSC San Diego)	
Seaweb Underwater Acoustic Nets	234
Joseph A. Rice, Robert K. Creber, Christopher L. Fletcher, Paul A. Baxley, Kenneth E. Rogers, and Donald C. Davison (SSC San Diego)	
Shallow-Water Acoustic Communications Channel Modeling Using Three-Dimensional Gaussian Beams	251
Paul A. Baxley, Homer Buckner, Vincent K. McDonald, and Joseph A. Rice (SSC San Diego)	
Michael B. Porter (SAIC/Scripps Institution of Oceanography)	
Advanced Refractive Effects Prediction System (AREPS)	262
Wayne L. Patterson (SSC San Diego)	
A Passive Ranging Technique for Objects within the Marine Surface Layer	267
Stephen Doss-Hammel (SSC San Diego)	
Silicon-on-Sapphire Technology: A Competitive Alternative for RF Systems	271
Isaac Lagnado and Paul R. de la Houssaye (SSC San Diego)	
S. J. Koester, R. Hammond, J. O. Chu, J. A. Ott, P. M. Mooney, L. Perraud, and K. A. Jenkins (IBM Research Division, T. J. Watson Research Center)	

Strategies for Optimizing Bandwidth Efficiency

Todd Landers
SSC San Diego

DATA CHARACTERISTICS

Loss-Sensitive Information

The data type envisioned when discussing data communications is usually loss-sensitive data. Many network engineers mistakenly assume that loss-sensitive data are the predominant data type on the Navy's general-purpose wide-area networks (WANs). This data type must be faithfully reproduced with 100% accuracy at the distant end of the link before it can be used. The Transmission Control Protocol (TCP)/Internet Protocol (IP) usually transmits the data type as computer data/software. Any differences in the data reproduced on the distant end of a link make the entire communication unusable. If data are corrupt, the application of the data will be invalid. The time it takes the data to reach the distant end of the link has little effect on data usability. This data type is loss intolerant, but latency tolerant.

As mentioned, the bulk of this data type is transferred using the IP family of protocols. IP is inherently connectionless. It uses a 32-bit address scheme to identify hosts on the network. The User Datagram Protocol (UDP) and TCP use IP as its transport layer.

For broadcast applications, UDP uses the connectionless properties of IP to its advantage. It uses a checksum to verify data integrity and discards corrupt data. UDP is ideal for applications that are loss and latency tolerant, and can be used in those instances to help minimize unneeded router chatter over bandwidth-constrained links.

TCP adds a connection-aware element on top of IP. TCP has embedded mechanisms that check the content and sequence of arriving packets. TCP also allows hosts to set timers. The host may "time out" a connection, enabling the host to free up system resources that would otherwise be tied up with a suspected dead connection. TCP will automatically request a re-send from the originating host if loss or corruption is detected. TCP uses these and other tools to keep the network working smoothly as long as latency is managed at lower levels of the network.

Unfortunately, the same features that help TCP work well in situations where bandwidth is ample can be disastrous when bandwidth becomes constrained. Once a link in the network becomes bandwidth-constrained, the applications start asking for retransmission of data assumed lost (in this case, just delayed). This unnecessary request for information is the

ABSTRACT

To optimize bandwidth efficiency, the natural limitations of each network-supported data type must be overcome or mitigated. This paper discusses issues affecting bandwidth efficiency through the U.S. Navy's bandwidth-constrained wide-area network (WAN). The paper details the prevalent data types found in the naval environment and describes the characteristics associated with each data type. Commercial, standards-based link layer protocols that have widespread application in Navy networks are also described. Finally, forward error correction and issues surrounding bandwidth efficiency are discussed.

beginning of the end of data transfer across the link. If Ethernet is used as the link layer protocol, connection timeouts caused by long round-trip time can cause a router to start seeking alternate paths to the desired host. The additional router chatter contributes to the congestion of the already congested link. This congestion is a death spiral for a TCP connection. The connection is terminated, and if the host is looking for the original information, it attempts to reconnect. Note that no useful information is exchanged, though bandwidth is consumed. The bandwidth consumption prevents other worthy circuits from exchanging useful information.

The way applications use IP may cause other inefficiencies. If the payload of the IP packet is not appropriately sized for the data type conveyed, huge amounts of bandwidth could be consumed because bit stuffing is needed to make complete packets. The data type usually originates at hosts that provide sensor inputs (like voice) to another application. If a sensor needs to transmit a sample containing a few bytes to a remote host through TCP, the originator usually stuffs filler bytes into a packet with a length that is probably several hundred bytes. The efficiency of this connection quickly approaches zero, which is not a problem until bandwidth becomes limited. File compression can reduce the size of an application data file or sensor output before transmission through the WAN. File compression engines must be deployed to all source and user sites, which causes a logistics problem, but the overall gain in bandwidth efficiency is worth the trouble. File compression can reduce the actual data transmitted across the link by 80%.

Header compression techniques can reduce bandwidth consumed over a point-to-point link through various network protocols. For header compression to add value to the WAN, it must add minimal latency due to processing overhead and be completely symmetrical. Compression abbreviates redundant header information in a data stream before transmission over the WAN and then restores the header to its original state after it reaches its final destination. The higher the compression engine is in the protocol stack, the more opportunity to save bandwidth. The more aggressive the compression engine, however, the more latency it adds to the circuit. Header compression at the transport and network layers depends on some error checking at the link layer to be effective.

Time-Sensitive, Loss-Tolerant Information (Voice)

The most common type of time-sensitive, loss-tolerant data is plain old telephone system (POTS). Unlike computer-oriented information, the interpretive device for POTS is the human ear. Studies show that over 50% of a speech signal can be removed and the human ear can still assemble the required information to extract the audible message. However, as little as a 0.5-second delay can cause severe degradation of the intended communication. A system designed to handle large quantities of this data type can lose a lot of data, but if data are delayed or delivered out of order, it is useless, and interpreted as noise.

Networks specializing in this data type are quite different from those that handle large quantities of loss-sensitive data. In terrestrial networks where bandwidth is ample, a typical voice call is digitized and transmitted using a G.711 protocol through the public switched telephone network (PSTN) at 64 kbps. Toll-quality voice has an upper latency limit of a 200-ms delay across the network. These networks are composed of various

sizes of public branch exchange (PBX) switches. The interconnections between PBXs generally scale in 64-Kbps chunks.

PSTNs are connection-oriented. When a call initiates, the originator transmits a setup preamble that negotiates for a connection at each intermediate switch along the way. If the connection cannot be supported at any point along the way, the entire connection is denied. If all attempts to establish the end-to-end connection are denied, the originator gets a busy signal. For most PSTN users, this busy signal only happens on Mother's Day or after a natural disaster. There is no such thing as a lower grade of service; a connection exists or it does not. During the call, a near-real-time connection for $N \times 64$ kbps allows the user to talk, send a facsimile (FAX), or use a modem, secure telephone unit (STU), or secure telephone equipment (STE), etc. After the call is completed, a teardown sequence allows each switch in the circuit to release the resources reserved for that connection.

This type of network has many sources of inefficiency. First, voice is the most common type of connection supported through this network. Voice typically has less than a 50% duty cycle. Generally, only one person talks at a time, and usually there is silence between words. Everything else is dead air (wasted bandwidth). Silence-suppression techniques reduce the dead-air bandwidth consumption to help solve this problem.

Another source of inefficiency is that 64 kbps is not really needed to digitize and communicate using voice. The 64-kbps convention was adopted because it was an easily implemented solution, not because it was the most efficient. Several voice compression algorithms can drastically reduce the amount of bandwidth used for each voice call. Toll-quality voice has been compressed to 8-kbps or one-eighth of the bandwidth allotted for a typical voice call. Good-quality voice has been transmitted using less than 800 bps. Unfortunately, other applications using the PSTN do not respond well when compressed with some of the more aggressive compression techniques, so compression must be applied selectively.

Modems and FAX machines use the PSTN to transmit analog-modulated digital signals. This inefficient means of digital data transfer was developed years ago to overcome noisy analog transmission lines that were once used to interconnect PBXs and end-users. Connections have improved, but the format is outdated. The most efficient way to accommodate these types of connections is to convert the modulated digital signal back into ones and zeros and transmit them through the network using much less bandwidth. A FAX machine can be supported at 9.6 to 14.4 kbps instead of consuming the full 64 kbps allocated to each connection by the PSTN. STU-III can also be supported using this type of compression technique, but the modulation scheme must be implemented in accordance with National Security Agency (NSA) policy.

Time and Loss-Sensitive Information

The synchronous serial data type is traditionally used where the system designer had creative control of the entire system. These links are susceptible to loss of content and fluctuations in the end-to-end timing. Each communication link was usually built to support one application set. Interoperability and flexibility were not considered in the design. These systems are probably the single largest source of wasted bandwidth. After the communication link initiates, it remains active regardless of use.

Circuits had to be provisioned to support worst-case bandwidth needs, and as applications became more bandwidth-efficient, their bandwidth usage remained high and constant.

As networked applications became more popular, synchronous serial communications became known as communications "stovepipes." The Department of Defense has invested huge amounts of resources into developing and refining stovepipe systems over the past 30 years. Although new systems focus more on networked solutions, stovepipes are still with us today primarily because of the cryptography developed to support legacy applications. The slow development of network cryptography has hindered application development and subsequent migration away from stovepipes.

Video

Video is generally more tolerant of timing than synchronous serial connections, but jitter is deadly. The type of video compression used should vary depending on the video content. Compressed video usually transmits all information needed to paint the screen the first time, and transmits only the changes to the initial image. Regardless of the resolution or quality of the video, this approach allows video to be supported using variable bit-rate service contracts through the network. Talking-head videoteleconference (VTC) video should use the most aggressive video compression techniques. This type of application can operate well on less than 64 kbps.

One of the worst misuses of bandwidth for video traffic occurs when the host or the network provisioning creates a fixed-bandwidth pipe for the video call. H.320 is a common video compression format used with Integrated Services Digital Network (ISDN) networks. Each video call allocates $N \times 64$ kbps to support the call resolution selected by the user. The bandwidth within the fixed allocation continues to fluctuate; however, even though the bandwidth need reduces when the picture stabilizes, no bandwidth is available for other applications.

Although still both time and loss sensitive, the H.323 protocol is much more tolerant in both areas. H.323 works with IP networks and has progressed in overcoming some of the inherent obstacles for supporting voice and video on an IP network.

Unfortunately, H.323 is susceptible to many problems that plague data transmission over IP-based WANs. While some jitter or delay can be tolerated, excessive congestion can cause the H.323 session to freeze. Though there is some time-sensitivity, packets are mixed with and sometimes delayed because of packets that have no time-sensitivity. To overcome this deficiency, priority queuing can reduce the likelihood that the time-sensitive video will incur fatal transmission delay.

The screen capture in Figure 1 shows the bandwidth consumption of the H.323 protocol generated using Microsoft Netmeeting. The link information shown represents one-half of a bidirectional link. The video resolution for this example was set up to run at best-fidelity voice and video. H.323 is highly variable in its bandwidth requirement (red trace in Figure 1). The

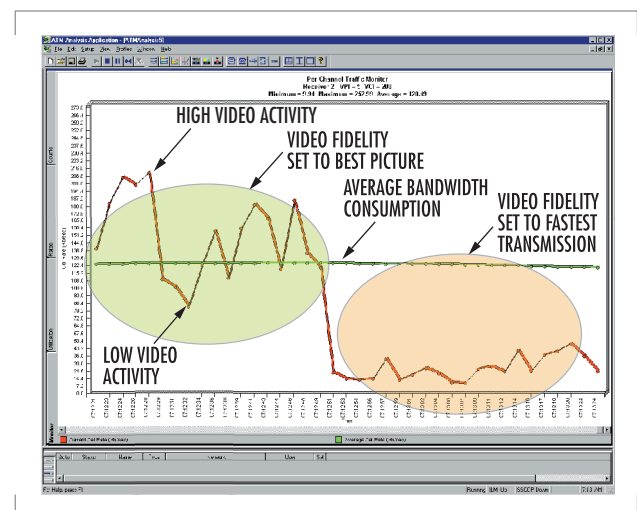


FIGURE 1. Bandwidth consumption of the H.323 protocol.

surges in bandwidth usage occur when the video compression engine must transmit updates to large portions of the image. With video quality set to its highest setting, the average bandwidth usage settles out near 130 kbps, with surges up to nearly 190 kbps. During periods when the video does not change, bandwidth usage drops to below 80 kbps. Netmeeting allows the user to reduce the fidelity of the picture to accommodate bandwidth-constrained connections. At the lowest resolution settings, the average bandwidth usage has been observed below 20 kbps for a full-motion VTC.

TYPICAL NAVY WAN DATA LOADING

The U.S. Navy operates in a truly converged WAN environment. Figure 2 shows a sample of the circuits assigned to USS *Coronado* (AGF 11), a command ship, during a typical deployment. The bandwidth available on *Coronado* should be considered the best possible case because she has been outfitted with the best communications available in the U.S. Navy to support various developmental enterprises.

Coronado runs multiple T-1s using various super high-frequency (SHF) and Challenge Athena configurations.

Voice, video, and data must all co-exist on the WAN (Figure 2). The major users include a Joint Service Imagery Processing System (JSIPS), which is an intelligence circuit currently using a synchronous serial EIA-530-based system. This circuit was one of the primary reasons for the procurement of the Challenge Athena system, so when this circuit becomes active, other lower priority circuits are manually disconnected. The JSIPS circuit is a prime example of current bandwidth management practices. Other large data users include secure and non-secure voice (both circuits are listed in the figure as POTS LINE). POTS lines are supported using compressed voice cards in various time-division multiplexers (TDMs). The compression cards reduce the bandwidth required to support each voice call from 64 kbps to between 8 and 16 kbps. The compressed voice signals are aggregated as synchronous serial circuits before porting to the satellite communications (SATCOM) modems.

In the bandwidth management approach, fixed-bandwidth synchronous serial circuits constrain circuits that use protocols that dynamically consume bandwidth such as IP. The circuits marked "ADNS" (Automated Digital Network System) represent the wide-area IP-based traffic, and typically are assigned up to 384 kbps, supporting a mixture of classified and unclassified data.

Figure 3 shows how the circuits on the Y-axis might consume bandwidth when provisioned through a TDM using today's provisioning approach. The white space represents provisioned, but unused, bandwidth. A few points to notice in this figure are as follows:

- Wasted bandwidth by low-usage, high-bandwidth systems such as the Video Information Exchange Subsystem (VIXS), which is a H.320-based VTC using synchronous serial cryptography for security

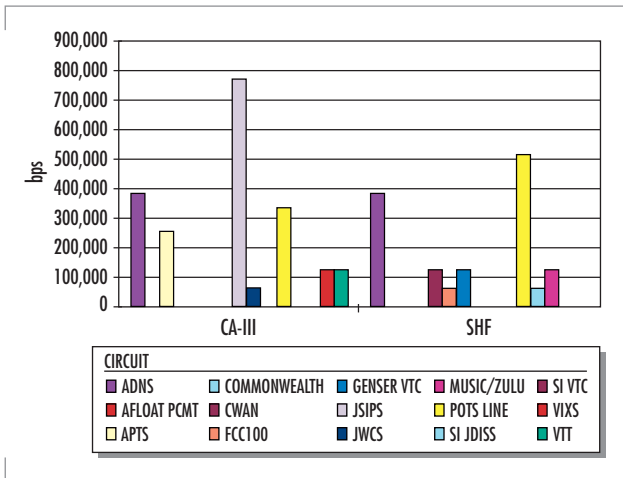


FIGURE 2. Circuits assigned to *Coronado* during typical deployment.

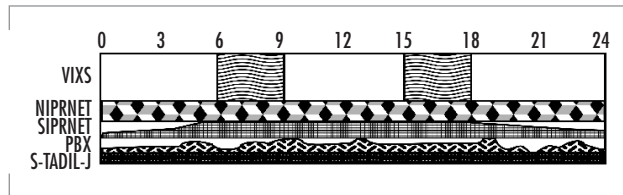


FIGURE 3. TDM circuit.

- Congestion in the low-priority, high-usage circuits such as NIPRNET (unclassified but sensitive IP router network)
- "Nailed up" circuits consuming bandwidth to keep the circuit timing alive, such as Satellite Tactical Data Information Link-Joint (S-TADIL-J)
- More bandwidth available to support additional voice and data circuits than are allowed under current provisioning policy

LINK-LAYER CONSIDERATIONS

Ethernet

Ethernet is commonly used as the link-layer protocol for TCP/IP and UDP/IP. Ethernet is a very cost-effective way to deliver data to end-users. The network equipment is inexpensive and mature; there is a large application base with drivers supporting Ethernet; and there is a large pool of competent network administrators who understand the technology. Ethernet is inexpensive to deploy and administer. It scales very easily to 10 Gbps on the backbone. Ethernet uses the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) approach to sharing bandwidth among the network users. If there is a collision, it simply re-sends the information. Ethernet works best on generously provisioned networks; for the most part, the assumption of few collisions is true.

Ethernet starts having difficulty when congestion occurs. At approximately 40% of the rated network throughput, Ethernet begins to bog down. At 60% link saturation, the link becomes nearly unusable because much of the traffic is re-sent from prior collisions.

Switched Ethernet technology provides answers to some of these problems by explicitly controlling traffic destined for users. An Ethernet switch can support single dedicated or multiple users on a switched segment. The switch logically separates the segments to eliminate collisions between segments. Multiple user segments continue to have the contention problem among subnetwork users.

Ethernet is problematic in the wide area because the organization deploying the local network will not have control over congestion in the wide area. Any single link in the wide area will slow performance experienced by the user. Network performance is only as good as the slowest link in the WAN.

Point-to-Point Protocol

Point-to-point protocol (PPP) is an encapsulation approach to transmitting IP over serial point-to-point links. PPP is a very flexible approach to transmitting IP datagrams through a serial link. The only real limitation PPP imposes on the link is that it has to be a full-duplex link. It supports synchronous and asynchronous transmission. It works fine over a number of common physical media including EIA 530, RS-232, V.35, etc. However, PPP links can cause unwanted latency and jitter because of the variable nature of the IP datagram contained in the data payload of the PPP frame.

ATM

Asynchronous transfer mode (ATM) effectively transmits a wide variety of data across a network. The size of the ATM cell (53 bytes) was developed as a compromise between the voice camp (small, prompt data

delivery) and the IP camp (large, continuous streams of guaranteed data delivery). While ATM was originally envisioned to work on high-speed networks (OC-3 and above), it has been adapted for the WAN because it works through congested links.

ATM statistically multiplexes fixed-size cells through a link. The protocol organizes cells into logical or virtual point-to-point circuits through an interface. At the time of circuit setup, each interface in the circuit establishes a service contract with its neighbors. Each switch has a unique address to ease automated connection setup. Once all of the interfaces in the path have established the required service contracts, the data transfer begins. Cells with that circuit identifier are automatically switched along its path to the end-user. Once the data transmission is complete, the contracts are canceled and the circuit is disestablished.

The service contracts have built-in quality-of-service features. At the top layer, there are ATM Adaptation Layers (AALs). Each AAL has some predetermined characteristics and some preconceived notions of what applications that adaptation layer would support. For example, AAL-1 supports synchronous serial connections and looks to users like a static TDM. AAL-2 supports voice, and while it maintains the time relationships between cells, it can take advantage of the other characteristics of voice discussed earlier. AAL-5 supports IP and has many features to take advantage of the characteristics of IP data transfer.

Forward Error Correction

Forward error correction (FEC) is commonly applied on noisy links to improve error performance and, thus, the performance of the link. The different FEC algorithms include block, convolution, and Viterbi codes. For the purposes of this paper, FEC can be applied in varying degrees to reduce the error rate of a link; however, the more rigorously FEC is applied, the more bandwidth overhead and processing latency increases.

As discussed, different data types have varying degrees of error tolerance. FEC should be tailored to the data type passing through the link. For example, voice is loss tolerant and will perform well even when the link has some errors. Some synchronous data streams are very loss sensitive and may cause the end-user equipment to malfunction if there are too many errors on the link. TCP/IP may start flooding the link with re-sends if the error rate is too high, thus causing data congestion.

Recently, many products have been shipping with adaptive FEC. This approach samples the noise on the link and adjusts the FEC algorithm to keep the link error rate nearly constant. As the FEC is applied more aggressively, the effective throughput drops. Applying adaptive FEC at the link layer provides better performance by reducing the amount of re-sends by the hosted applications.

Link Design Decisions

As with all aspects of an engineered solution, engineers must choose the best tools to confront each aspect of the link design. Because of the economics and maturity of the technology, Ethernet is a clear choice in the local area networks, but falls short in the wide area. PPP is a good choice if all of the applications supported by the network are IP-based, but PPP falls short for a general-purpose network that supports various data types. With the technology currently available, ATM is the only technology

discussed that can support a truly converged network supporting voice, video, data, and legacy applications. Figure 4 shows the dynamic bandwidth allocation achieved using ATM for the WAN.

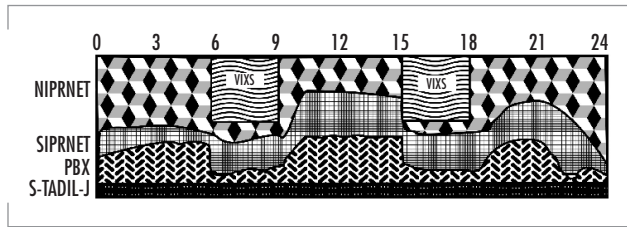


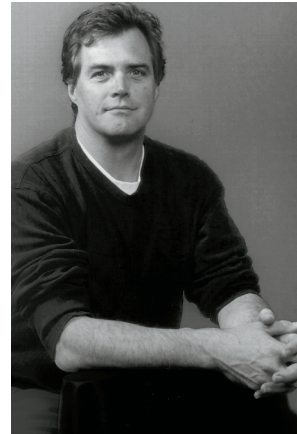
FIGURE 4. ATM circuit.

CONCLUSION

We can objectively maximize the information payload, optimizing for communication channel size, reducing non-productive information transfer, and using aggressive compression and forward error correction techniques. To get the greatest use from a SATCOM system, we must maximize the gross bandwidth efficiency. Additionally, user application data must be optimized. TCP/IP or UDP/IP solutions should be implemented wherever possible. Finally, the correct link layer technology must be selected for the environment in which it will be deployed.

Successful information transfer occurs only when enough data are transferred from the source through a data link to the distant end of the link to facilitate reassembly of the information suitable for end-device perception. Various types of information will be transmitted through our communications links. Some information is loss sensitive; other types of data are time sensitive; and still others are time and loss sensitive. They are all vital to the operation of the Fleet.

There are opportunities at every layer to optimize. The fiscal cost of not optimizing is tremendous. The operational cost could be devastating.



Todd Landers

BS in Electrical Engineering,
San Diego State University,
1985

Current Research: Wide-band
ADNS architecture develop-
ment; Tactical Switch System
development; DoD Teleport
requirements analysis and system
definition; design/test of the
IT-21 block one baseband
system.

Tools for Analyzing and Describing the Impact of Superstructure Blockage on Availability in Shipboard and Submarine Satellite Communications Systems

Roy A. Axford, Jr.
SSC San Diego

Gerald B. Fitzgerald
The MITRE Corporation

INTRODUCTION

On most of today's warships, it is impossible to find a single location for a satellite communications (SATCOM) antenna that provides an unobstructed view of the entire sky. If there is sufficient available topside space, two antennas are usually installed to support a mission-critical system (e.g., protected extremely high frequency [EHF] SATCOM). Lower priority systems are often forced to use a single antenna.

No matter how many antennas are used, it is critical to quantify the impact of topside blockage on the availability of a shipboard SATCOM system. Knowledge of this impact is needed in antenna location selection to ensure that the highest priority systems have the best views of the sky. Presenting this knowledge in an easily understood manner can make the topside design process more successful. Furthermore, once a shipboard antenna system is installed, the ship's company must have a clear understanding of the impact of unavoidable blockage on communications availability.

Many ship captains say that there are times when their choice of heading is dictated by whether or not a particular antenna system can "see" a desired satellite. It also appears that the determination of unblocked headings is often made by trial and error at sea, without benefit of *a priori* knowledge of the blockage situation of the SATCOM terminal in question. Topside blockage is so frequently discussed in the Fleet that there is widespread need for a software tool that can present this knowledge clearly.

This paper describes a set of tools for the analysis and display of the impact of superstructure blockage on shipboard SATCOM availability. These tools can give a ship's crew real-time indications of the blockage situation for an antenna system of interest with any desired geostationary satellite, based on the ship's present position, heading, and the Sea State. (Geosynchronous satellites in inclined orbits are discussed later in the Spatial Model section.) For route planning, there is also a display that shows the blockage situation along an entire Great Circle path as a function of Sea State (for the Great Circle headings). For more general planning and analysis, there is a display that shows the percentage of blockage-free headings (as a function of position and Sea State) as a colored cell on a global map. Along with their value to the operational community, these tools help topside designers compare the relative merits of candidate antenna installation locations.

ABSTRACT

The blockage analysis tools in the satellite communications (SATCOM) Availability Analyst (SA2) software package combine topside blockage data, communications satellite constellation positions, and ship-motion models to calculate the impact of superstructure blockage on availability as a function of antenna installation locations, ship's geographic position, and sea state. This impact can be evaluated for a specific position, along a ship's planned route, or averaged across the entire field of regard of the SATCOM constellation(s) of interest. This paper details the capabilities of the blockage analysis tools in SA2. The tools are applied to the analyses of topics of current interest including International Maritime Satellite (INMARSAT) on the CG 47 class, Global Broadcast Service (GBS) aboard the flagship USS Coronado (AGF 11), and submarine High Data Rate (SubHDR) on the SSN 688 class.

The blockage analysis and display tools described here are components of a larger program called SATCOM Availability Analyst (SA2). The next section, Component Models, describes the lower-level components used to model the effects of blockage in SA2 (i.e., inputs). The Displays and Metrics section presents SA2's blockage analysis products (i.e., outputs). The Applications section gives some illustrative examples of SA2's recent application to the analysis of blockage for emerging SATCOM terminals to be installed on Aegis cruisers and *Los Angeles* class submarines.

COMPONENT MODELS

SA2 was developed as an extension of the Global Broadcast Service (GBS) Data Mapper (GDM) [1, 2]. GDM combined a simple Java-based Geographic Information System (GIS) with encodings of relevant International Telecommunications Union Radiocommunications (ITU-R) Recommendations and GBS link budget parameters to develop global maps of GBS link margin and availability.

Spatial Model

The core of SA2 is this same GIS, built upon a simple raster model of the earth's surface, which is represented as an array of $2.5^\circ \times 2.5^\circ$ *model cells* (between latitudes 70S and 70N). The model-cell center points are stored as 3-space (x, y, z) vectors. The model-cell size can be varied, but 2.5° represents a good trade-off between precision and run time for most applications.

SATCOM constellations in SA2 are also represented as sets of 3-space vectors—each vector giving the Clarke Belt position (CBP) of one geosynchronous satellite. SA2 computes the elevation and azimuth angles from the center point of any model cell to a satellite's CBP by using simple vector arithmetic. These angles may then be combined with a ship's heading (entered by the SA2 user), and used as indices into a blockage matrix. As described in the following section, this matrix is an image of the superstructure blockage as seen from each antenna assigned to the satellite of interest. Thus, SA2 computes whether or not a ship in a given model cell and on a particular heading has an unblocked line of sight (LOS) to the satellite of interest. In addition, the SA2 user may enter a Sea State. SA2 then uses ship-class-dependent motion models (described later in the section on Ship Motion Models) to expand the LOS from the antenna into an appropriately distorted cone, thus accounting for the impact on satellite visibility. Increased ship motion in higher Sea States reduces availability by causing the superstructure to move in and out of an antenna's LOS to the satellite of interest. As shown in examples below, some antenna locations suffer more from this effect than others.

Many geosynchronous communications satellites of interest are in inclined orbits (e.g., the Ultra-High-Frequency [UHF] Follow-On [UFO]/Global Broadcast Service satellites: UFOs 8, 9, and 10). The pointing angles to such satellites from a geographic position vary over the diurnal cycle. SA2 does not model this motion, but it is able to read satellite track files in the form of pairs of azimuth and elevation pointing angles versus time for the position and satellite of interest. Such track files are readily available from applications such as Satellite Tool Kit (STK) and Satellite Orbit Analysis Program (SOAP). SA2 can use these pointing angles similarly to those for geostationary satellites (i.e., a geosynchronous

satellite in an orbit with 0° of inclination with respect to the equatorial plane) to determine if the LOS is unblocked.

Blockage Models

Only the highest antenna on a ship can view the entire hemisphere of sky above it, and even then, only if the antenna is also the highest structure of any kind on the ship. Otherwise, additional antennas, masts, exhaust stacks, weapons, yardarms, or any other superstructure will mask out (i.e., block) some of the sky. The cluttered view of the sky from a shipboard SATCOM antenna's topside location is represented in SA2 as a two-dimensional, binary-valued matrix, with 360 columns covering azimuth angles in 1° increments, and 106 rows, covering elevation angles from -15° to the zenith (90°). (The zenith row is a degenerate case; all entries are identical.) Depression angles (elevation angles below 0°) must be included because as a ship rolls and pitches, the apparent elevation angle to a satellite near the horizon (relative to the ship's deck) may be negative.

Blockage matrices may be imported into SA2 by two methods. For installed terminals, the blockage information often already exists in a Blockage Adaptation Module (BAM) file generated from a digital image(s) of the view(s) from the antenna location(s). Alternatively, direct processing of such a digital image can create a blockage matrix for SA2. A suitable image could be acquired from the antenna's installation location with a fisheye-lens-equipped camera, but this method has not been used to obtain any of SA2's blockage data thus far. Almost all of the blockage data used in SA2 come from images generated by a three-dimensional computer-aided design (CAD) model of the ship's entire topside. Such topside models are often refined and/or updated by taking theodolite surveys of the ship's topside directly from the intended antenna installation location(s). (Reliance only on a ship's design drawings can lead to the omission of superstructure that was added after initial construction.) The image-processing software used is external to SA2.

The software can acquire and digitize images in either polar or rectangular projections, and the software is an extension of a MITRE-developed image processing and exploitation suite originally written for the National Imagery and Mapping Agency (NIMA). Figure 1 presents a CAD-model topside blockage image typical of those that have supplied most of the blockage matrices available within SA2.

Note that these matrices model boresight or optical blockage. SA2 does not consider the near-field patterns of shipboard antennas or effects such as knife-edge diffraction. (This approach is supported for frequencies above ~ 1 GHz by conclusions of a study [3] in which detailed tests and analyses were performed to determine the blockage effects of various topside structures on the performance of the AN/USC-38 EHF shipboard SATCOM terminal.) However, antenna beamwidth can be simulated within SA2 by a simple, run-time operation that pads each blocked area by a user-specified number of degrees. A similar procedure is often used in producing the BAM files of

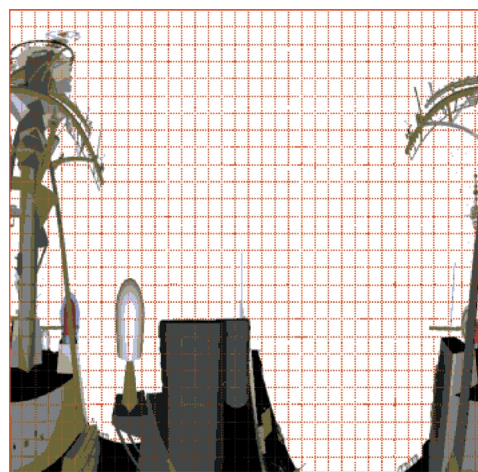


FIGURE 1. 3-D CAD topside model blockage image (rectangular projection) of the view from one of the INMARSAT antenna locations on the DDG 51 class.

dual-antenna systems to mark a "warning track" for the initiation of antenna handover procedures. Furthermore, BAM files often designate some of the smaller (say, less than 5 to 10 degrees in azimuth or elevation extent) unblocked areas as "blocked" to avoid "peep holes" that are lost due to ship's motion in moderate Sea States. This practice has also been adopted in producing the blockage matrices used SA2.

For dual-antenna systems, SA2 determines when an unblocked LOS to a desired satellite is available from *either* antenna or from *both* antennas. The "either antenna" mode analyzes systems that perform a hand-over from antenna "A" to antenna "B" as "A" moves into a "warning track." In these systems, either antenna can provide 100% of the communications services if it has an unblocked LOS to the satellite. The "both antennas" mode analyzes systems such as INMARSAT B High Speed Data (HSD) in which both antennas track the same satellite simultaneously to provide higher total throughput by using multiple transponder channels. In these systems, both antennas are required to provide 100% of the communications services.

Ship-Motion Models

As ships are accelerated by the wind and waves through which they travel, they experience Sea-State-dependent perturbations that are described by three rotational motions (pitch, yaw, and roll) and by three translational motions (surge, sway, and heave). These effects are detailed in [4]. Following McDonald's ranking of the magnitudes of these motions, SA2 confines itself to the impact of pitch and roll. Sea States high enough to make the other motions significant with respect to blockage are so severe that they surpass the operational specifications of Navy shipboard SATCOM antennas.

In [4], McDonald provides tables of length at the waterline, beam at the waterline, metacentric height and roll constants for various surface-ship classes. McDonald combines these constants with the ship-motion equations of DoD-STD-1399-301A [5] to provide ship-class-dependent pitch and roll extremes and periods as functions of Sea State. The resulting sinusoidal ship-motion equations are used to produce pitch and roll angles as functions of time for an animated display in SA2 in which the observer's frame of reference is the ship. These equations also allow the computation of temporal statistics (e.g., unblocked time/blocked time, durations of blocked times, etc.) that are of potential value in the evaluation of protocols for intermittent links.

Motion data are not as readily available for submarines. Since a helmsman actively controls the pitch of a submarine at periscope depth by using the stern planes, pitch is not a key factor in determining LOS availability at moderate Sea States. Roll is important, however. Submarine roll rates (even more than surface-ship roll rates), depend on the heading of the boat relative to the swell direction. For analyses of blockage aboard submarines, we have relied on interviews with former submariners to characterize the expected pitch-and-roll extremes and periods.

DISPLAYS AND METRICS

The data from the models described above are used in SA2 to provide blockage information for a ship's position along a route of travel or averaged over the entire field of regard of a satellite constellation of interest.

Ship's Position Blockage Information

SA2 provides an extremely useful display to assess the availability of a given SATCOM system from the ship's current position (or any position of interest) as determined by superstructure blockage and Sea State. Figure 2 provides three examples of SA2's SkyView display [6]. The currently selected blockage matrix (e.g., for a single antenna or for the composite blockage of two antennas) is always displayed in a polar orthographic projection. Ship's position may be typed in or entered by clicking on the desired spot on SA2's map display (see Figure 3). Ship's heading may also be typed in or adjusted with a slider. The satellites of the selected constellation (those above the horizon for the entered position) are then plotted as an overlay at the azimuth and elevation pointing angles computed for calm seas (i.e., for a level ship). Blocked satellites are shown in red, and visible ones are shown in green.

As noted above, the user can also examine the effects of Sea State with the SkyView display. Sliders allow the user to enter static pitch-and-roll angles based on, for example, the way the ship is behaving while underway (or to account for a list at the pier). The satellite's "dot" moves accordingly and turns red if it moves into a blocked region. Alternatively, using the pitch-and-roll magnitudes from [4], SA2 will plot the entire ship's motion envelope for a user-entered Sea State, resulting in green, red, or green and red "satellite smears" over the extent of pitch and roll. (Obviously, the user must correctly enter the ship's class for this approach to be useful.) The full equations of motion (according to the ship's class) may also provide an enlightening real-time animation of the apparent satellite positions. For example, see Figure 2 and imagine the satellite position moving according to the ship's equations of motion. (It is actually possible to get seasick while watching this display!) As the satellite moves, it turns red when blocked and green when unblocked. With all of these approaches, if any red appears, it indicates that the ship's motion might be causing intermittent outages for the SATCOM system in question. Thus, the SA2 SkyView display provides an aid for troubleshooting at sea.

Ship's Route Blockage Information

The SkyView display can help analyze SATCOM availability at a moderate number of positions, but availability along an entire ship's route is best viewed on SA2's TrackView display. With the aid of a text editor, the user enters pairs of end-points that are then connected by SA2 through use of a Great Circle route. The resulting tracks are color-coded along their extent according to the availability of the currently selected satellite or constellation of satellites.

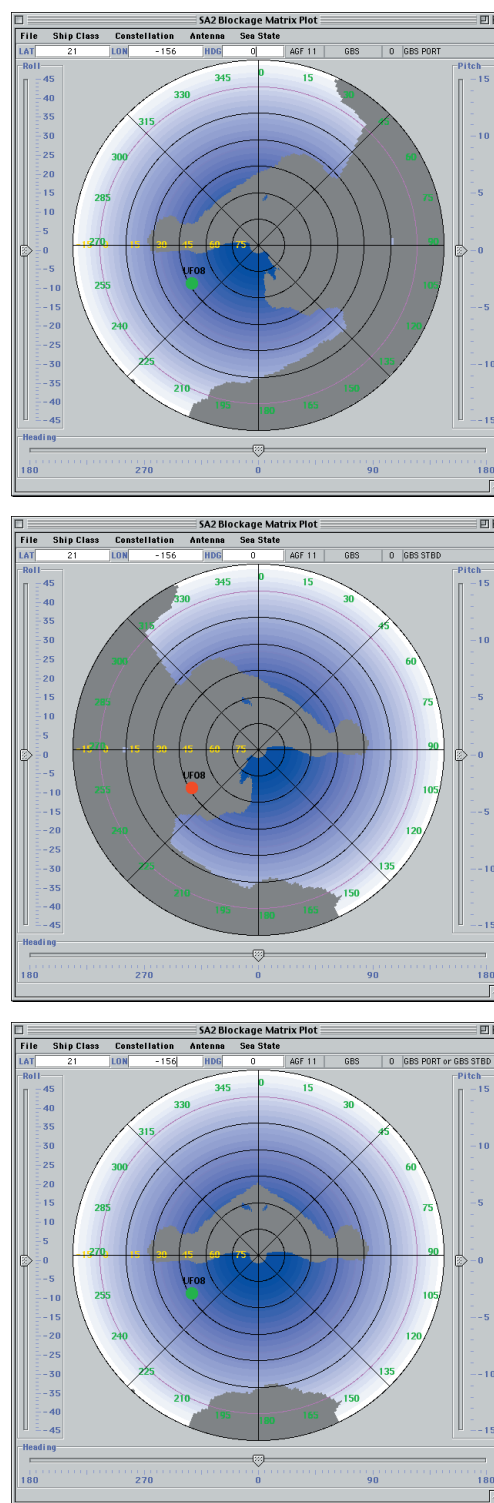


FIGURE 2. SA2 SkyView display for USS *Coronado*'s original GBS antenna installation locations. Top: port antenna. Middle: starboard antenna. Bottom: composite blockage for both antennas in "either antenna" mode.

The TrackView display can reflect different Sea States. It may also be animated at an accelerated speed of advance. In this animated mode, the SkyView display is slaved to the TrackView and updates as the track is traversed.

Figure 3 shows the TrackView display of the availability of the GBS due to blockage aboard USS *Coronado* (AGF 11) (corresponding to the blockage matrix in the bottom of Figure 2) in Sea State 4 along Great Circle routes between San Diego, CA, and Pearl Harbor, HI, and Yokosuka, Japan [6]. As with all SA2 TrackView displays, the blockage information plotted in Figure 3 assumes that the ship remains on Great Circle headings. With reference to the bottom of Figure 2, it is obvious that a significant number of *Coronado's* headings are blocked for GBS when the satellite in use appears above 30° elevation. This is most clearly seen using SA2's global blockage statistics and the Average Line-of-Sight Availability (ALA)View discussed in the next section.

Global Blockage Metrics

Normally, ships do not always maintain Great Circle headings while they are at sea. In general, a ship could be on any heading at a given moment depending on mission demands (e.g., flight quarters, zigzagging, etc.). Therefore, in considering the relative merits of alternative antenna installation locations, an important metric is *the percentage of headings that yield an unblocked LOS to the satellite of interest*. The following describes how SA2 calculates and displays this metric.

For any given Sea State, SA2 determines, for a combination of (1) satellite of interest, (2) ship antenna's location (or antennas' locations), and (3) ship's position and heading, whether the antenna(s) has (have) an unblocked LOS to the satellite of interest throughout the resulting ship-motion envelope. If, for a given position and heading, the satellite is visible throughout the entire ship-motion envelope, then that position is considered unblocked on that heading in the selected Sea State for the desired satellite. In performing this evaluation, by default, SA2 considers the heading blocked if the satellite is blocked at any point in the ship's motion envelope for the selected Sea State. This criterion is realistic for any bulk-encrypted link in which the encryption devices must "not miss a beat" to maintain synchronization. However, this criterion can be modified for alternative studies.

For each model cell, LOS availability is computed in the manner described in the preceding paragraph for all headings in 1° increments. The number of unblocked headings, divided by 360 and expressed as a percentage, is the ALA metric, a new figure-of-merit for analyzing blockage introduced in [6]. SA2 repeats this process for all spatial model cells. The resulting array of percentages is displayed as a color-coded map,

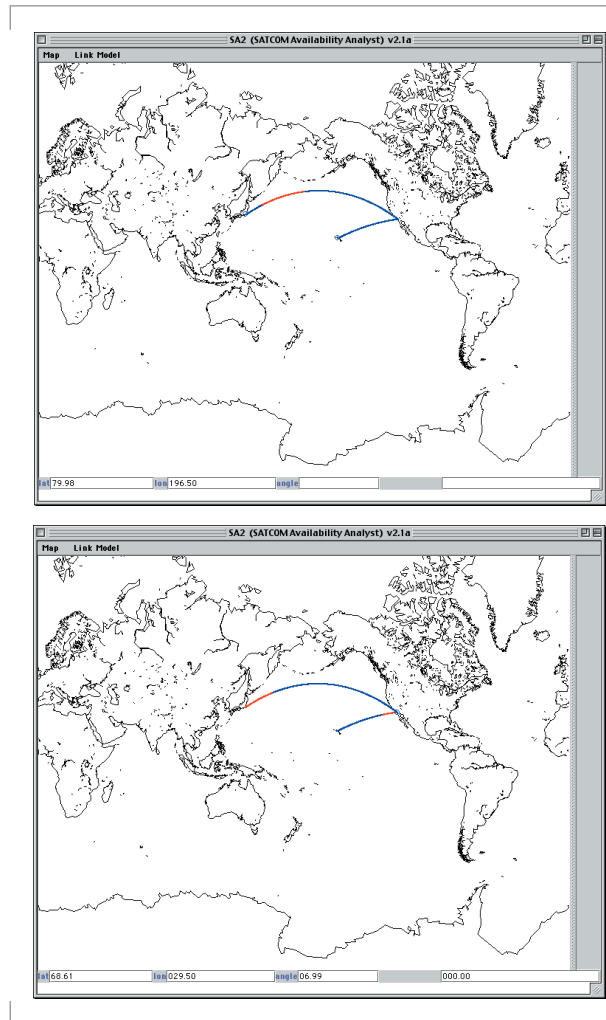


FIGURE 3. SA2 TrackView display for USS *Coronado's* original GBS antenna installation locations. Red: blocked; Blue: unblocked. Satellite in use: UFO 8. Sea State 4. Top: routes outbound from San Diego. Bottom: routes inbound to San Diego.

which is SA2's ALAView. These maps show, at a glance, where a particular shipboard SATCOM terminal is unblocked at all headings, at some headings, or at no headings. Figures 4A and 4B show ALAViews for the original GBS antenna installation locations aboard *Coronado*, in calm seas, and in Sea State 6, respectively, with the GBS transponders on Ultra-High-Frequency Follow-On (UFO) satellites 8, 9, and 10 [6].

The array of ALA figures spanning the satellite constellation's field-of-regard can also be averaged, yielding a *Global ALA* (GALA) metric. GALA can be calculated in two ways: (1) the average ALA over only ocean and littoral model cells within the field-of-regard or (2) the average ALA over all model cells within field-of-regard. SA2 uses the first definition by default.

APPLICATIONS

SA2 has been used recently to analyze INMARSAT B HSD availability aboard *Ticonderoga* class (CG 47) cruisers [7], submarine HDR (SubHDR-GBS and EHF) availability aboard *Los Angeles* class (SSN 688) submarines [8], and GBS availability aboard *Coronado* [6]. A detailed account of the *Coronado* work is reported in [6]. This section summarizes some of the conclusions of the CG 47 and SSN 688 analyses.

INMARSAT B HSD on the CG 47 Class

A commercial off-the-shelf (COTS) INMARSAT B HSD shipboard terminal has a single antenna and can support up to 64 kbps. To achieve 128 kbps throughput to the CG 47 and DDG 51 classes, it has been proposed to outfit each ship with two complete INMARSAT B HSD terminals. Additional INMARSAT space segment resources would be leased so that each ship would have access to an aggregate of 128 kbps by "summing" the 64 kbps channels from each terminal. Each COTS

INMARSAT B HSD terminal is an independent single-antenna system. There is no tracking hand-off from one terminal's antenna to the other. Therefore, to maintain a 128-kbps aggregate, each of the two antennas must be able to view the satellite continuously as the ship maneuvers. An analysis of the impact of blockage on the availabilities that this setup would achieve was accomplished using SA2's blockage tools in "both antennas" mode (see section on Blockage Models). For comparison, "either antenna" mode was also used.

Figure 5 shows ALAViews and GALA values for Sea States 0 and 6, in INMARSAT B HSD "both antennas" mode (128 kbps) and (a hypothetical) "either antenna" mode with handover (64 kbps) using the CG 47 class INMARSAT antenna installation locations. These results clearly show

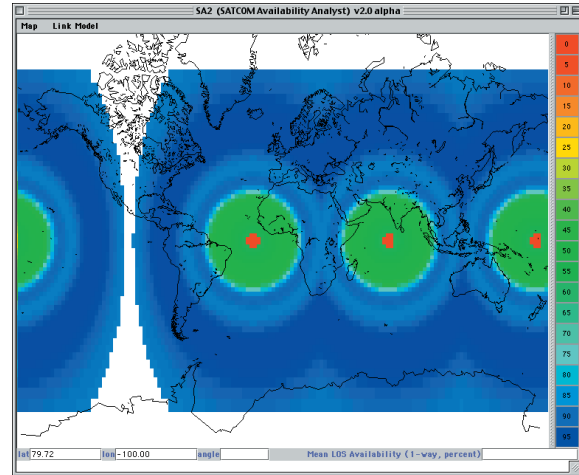


FIGURE 4A. ALAView for GBS aboard USS *Coronado*, original antenna installation locations, Sea State 0. GALA = 83.7%.

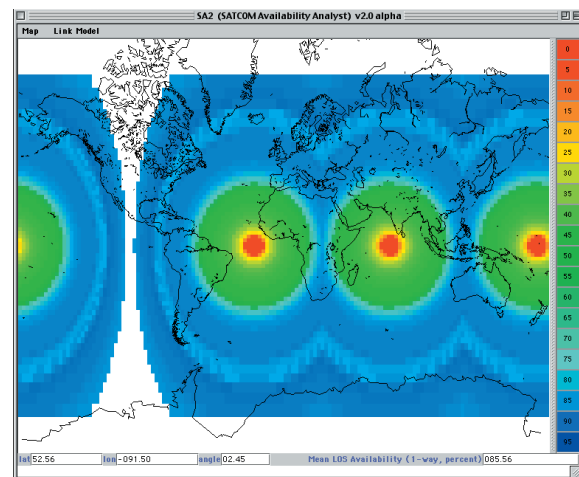


FIGURE 4B. ALAView for GBS aboard USS *Coronado*, original antenna installation locations, Sea State 6. GALA = 70.9%.

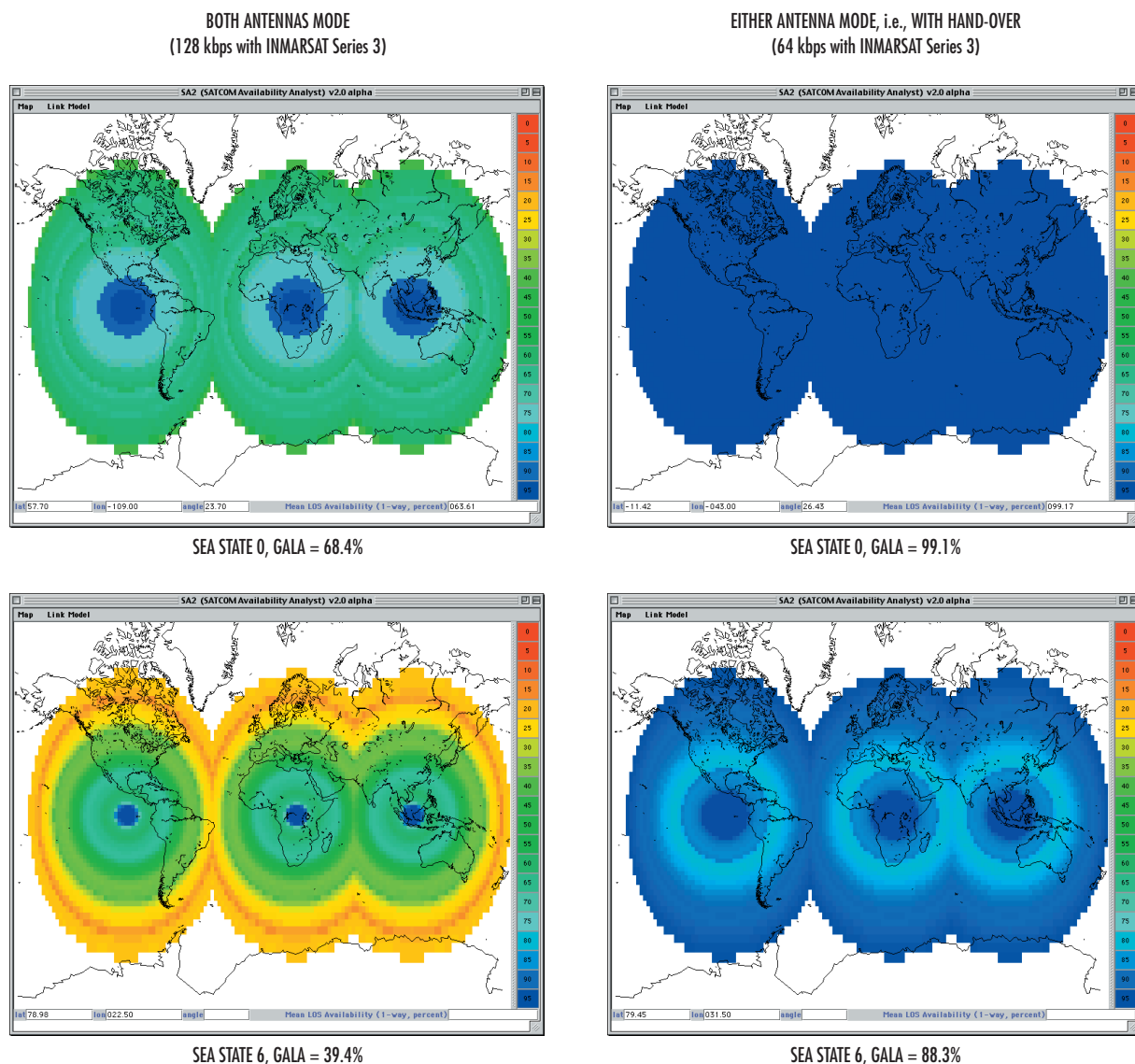


FIGURE 5. Dual-antenna INMARSAT availability on the CG 47 class analyzed with SA2's ALAView. Note: Only the three INMARSAT satellites on which leased services were offered at the time of the analysis are shown.

that a second, parallel INMARSAT B HSD terminal provides somewhat limited availability of a 128-kbps aggregate for the CG 47 class. However, the results also show that a dual-antenna INMARSAT terminal that would accomplish hand-overs between the two antennas provides an outstanding availability of 64 kbps, even in Sea State 6. After considering these results in July 2000, the Space and Naval Warfare Systems Command (SPAWAR) decided to investigate the development and acquisition of a handover-capable, dual-antenna INMARSAT B HSD terminal. Note that the next series of INMARSAT satellites, Series 4, will provide single-channel data rates up to 400 kbps, potentially making a handover-capable, dual-antenna INMARSAT terminal an even more valuable asset.

SubHDR on the SSN 688 Class

It is perhaps initially surprising that blockage is an issue for submarines, since the topside environment would appear to have no obstructions. In fact, Figure 6 shows there are several structures in close proximity to one another on the sail of the SSN 688 class. The short distances between them causes each to subtend a large solid angle as seen by the others. Furthermore, the masts and periscopes can be raised or lowered independently to variable heights.

The SubHDR system brings multiband SATCOM to submarines, including enhanced EHF capabilities and GBS. Early sea trials of the SubHDR mast and antenna system aboard USS *Providence* (SSN 719) revealed that from positions in the North Atlantic, the LOS to UFO 9 was sometimes blocked.

For analyses of SubHDR availability, SA2 represents the periscopes and other masts independently, each as seen from the point of view of the SubHDR antenna. Thus, the number of possible blockage matrices is large, but not all of them are tactically significant. For example, by doctrine, when a submarine is at periscope depth and any mast is raised above the waterline, a periscope must also be raised. Figures 7 and 8 present examples of SubHDR blockage matrices, which correspond to the first and sixth rows of Table 1. In all cases, the SubHDR mast is lowered 14 inches from its maximum possible height to avoid blocking the periscope. Table 1 shows GALA figures for six cases of equipment raised in addition to the SubHDR mast. In Sea State 3, the SubHDR GALA figure for the GBS payloads on UFOs 8, 9, and 10, or for the EHF LDR payloads on the same spacecraft, is, at best, 90% if the Type 8 Mod 3 periscope is used and 85.7% if the Type 18 is used. Figure 9 shows an ALAView for SubHDR, assuming that only the Type 18 periscope is raised.

Clearly, it is possible to analyze blockage for submarines with SA2 by using the same tools employed for surface ships. However, at sea, blockage is a somewhat different issue for submarines than for surface ships because submariners are generally more at liberty to select blockage-free headings after reaching periscope depth (PD). For example, submariners are never concerned about orientation with respect to wind direction in order to launch or recover aircraft. Furthermore, submariners often do not stay at PD

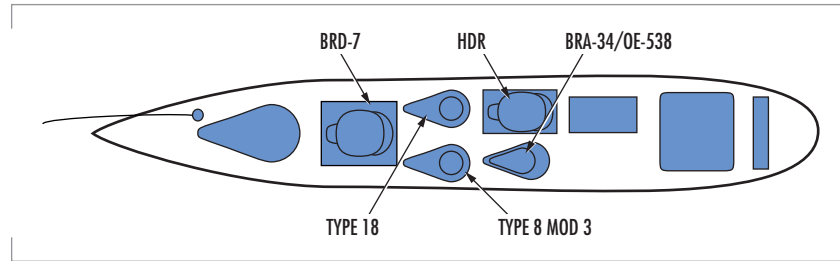


FIGURE 6. Sail configuration for USS *Providence* (adapted from [9]). The Types 8 Mod 3 and 18 are periscopes. The BRD-7 and BRA-34/OE-538 are multi-purpose masts.

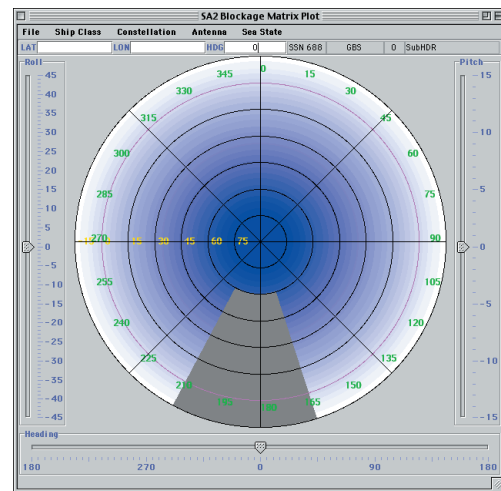


FIGURE 7. SubHDR blockage matrix when only the Type 18 periscope is raised.

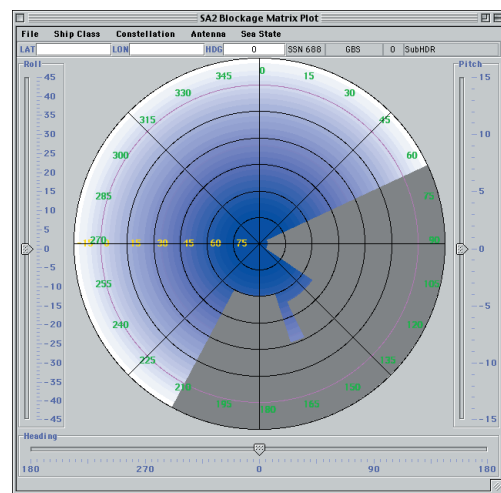


FIGURE 8. SubHDR blockage matrix when both periscopes are raised (Types 8 Mod 3 and 18) as well as both the BRA-34 and BRD-7.

any longer than is necessary to send and receive a few queues of communications traffic. On the other hand, in rough seas at periscope depth, submariners prefer to select headings more or less directly into the swells in order to minimize roll. In any event, SA2's SkyView display is useful to submariners for selecting blockage-free headings when using SubHDR.

SUMMARY

SA2 combines a set of simple mathematical models of the earth and of satellite constellations, coupled with similarly straightforward models of ship motion and of superstructure blockage to produce a powerful tool for assessing the impact of blockage on shipboard and submarine SATCOM availability. All of these components were previously available in various forms, but they had never before, to our knowledge, been combined in a single, simple-to-use package.

This paper has shown that various metrics are necessary to fully describe the impact of superstructure blockage on SATCOM availability over the full set of conditions in which ships and submarines serve. We also believe that this paper and our experiences using SA2 to interact with personnel from the operational, acquisition, and RDT&E communities have demonstrated that colored graphical displays are not just desirable, but are necessary to fully convey the impact of blockage on SATCOM availability.

REFERENCES

1. Fitzgerald, G. and G. Bostrom. 1999. "GBS Data Mapper: Modeling Worldwide Availability of Ka-Band Links Using ITU Weather Data," *Proceedings of the IEEE Military Communications Conference (MILCOM '99)*, http://www.agreenhouse.com/society/TacCom/papers99/48_3.pdf
2. Fitzgerald, G. and G. Bostrom. 2000. "GBS Data Mapper: Modeling Worldwide Availability of Ka-Band Links Using ITU Weather Data," *Proceedings of the 6th Ka-Band Utilization Conference*, pp. 217–224.
3. Brown, E. (RF Microsystems). 1993. "Navy EHF Program (NESP) AN/USC-38(V) Antenna Blockage Characterization," Naval Command, Control and Ocean Surveillance Center, RDT&E Division (NRaD),* San Diego, CA, (12 June), (available in the SSC San Diego EHF SATCOM In-Service Engineering Agent [ISEA] Library, catalogued as document 5000-362-19071).
4. McDonald, M. 1993. "SHF SATCOM Terminal Ship-Motion Study," TR 1578 (March), Naval Command, Control and Ocean Surveillance Center, RDT&E Division (NRaD),* San Diego, CA.
5. Department of Defense. 1986. "Interface Standard for Shipboard Systems, Ship Motion and Attitude," DoD Standard 1399, Section 301-A, (July), Washington, DC.

*now SSC San Diego

TABLE 1. GALA figures for GBS via SubHDR aboard the SSN 688 class for various combinations of equipment raised. The combinations represented by italicized rows, while technically possible, are not allowed by submarine doctrine.

Equipment(s) Raised (in addition to SubHDR)	GALA (%)	
	Sea State 0	Sea State 3
Type 18 Periscope	89.4	85.7
<i>BRA-34 Mast</i>	<i>83.7</i>	<i>79.5</i>
Type 8 Mod 3 Periscope	93.2	90.0
<i>BRD-7 Mast</i>	<i>96.4</i>	<i>94.9</i>
Type 18 and BRA-34	73.1	65.2
Types 8 Mod 3 and 18 plus BRA-34 and BRD-7	66.0	59.1

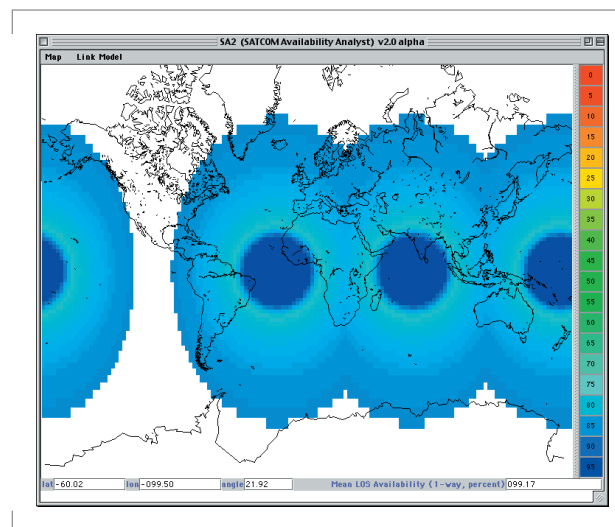


FIGURE 9. ALAView for GBS via SubHDR aboard the SSN 688 class. Type 18 periscope is raised; SubHDR mast is lowered 14 inches from its maximum possible height. Sea State 3. GALA = 85.7%.

6. Axford, R. and G. Fitzgerald. 2000. "Global Broadcast Service (GBS) Blockage Assessment for USS *Coronado* (AGF 11)," TR 1842 (November), SSC San Diego, San Diego, CA.
7. Colvin, B. 2000. "INMARSAT HSD on the CG 47 and DDG 51 Classes," (31 July presentation), SSC San Diego, San Diego, CA.*
8. Fitzgerald, G. 2001. "GBS Blockage Analysis for USS *Providence*," SubHDR Test Plan Working Group (TPWG), Session 8, (23 January), Space and Naval Warfare Systems Command (SPAWAR), PMW 173 (Submarine Communications), San Diego, CA.*
9. Chief of Naval Operations. 1998. "Submarine Communications Master Plan," (April), Washington DC, p. 4-2.



Roy A. Axford, Jr.

Ph.D. in Electrical Engineering,
Communications Theory, and
Systems, University of
California at San Diego, 1995
Current Research: Technologies
for wideband mobile satellite
communications.

Gerald B. Fitzgerald

BA in Linguistics and Computer
Science, Yale, 1977
Current Research: RF propaga-
tion modeling; imagery and
SIGINT fusion; network
intrusion detection.

*For further information, contact author.

Advanced Enclosed Mast/Sensor (AEM/S) System

John H. Meloling
SSC San Diego

ABSTRACT

The Advanced Enclosed Mast/Sensor (AEM/S) System is a revolutionary advancement in the topside design of Navy ships. Constructed of advanced composites, the AEM/S System is a self-supporting enclosed mast structure that provides affordable radar signature control and improved shipboard antenna system performance.

INTRODUCTION

The Advanced Enclosed Mast/Sensor (AEM/S) System uses advanced composites to produce a mast structure that encloses the existing legacy antenna systems of the ship. This enclosure consists of a composite sandwich structure that supports all internal decks, antennas, and ballistic cable trunks. Embedded within the composite sandwich are frequency selective surface (FSS) layers that filter electromagnetic waves. This filtering allows transmission and reception at desired frequencies while rejecting threat radar signals. Once these electromagnetic characteristics are designed into the composite sandwich, the mast structure can be shaped to reduce the radar cross section (RCS).

The AEM/S technology has many advantages. AEM/S provides affordable signature control of legacy antenna systems. Developing and fielding new antenna systems is a long and costly process. The AEM/S System provides a near-term means of reducing the RCS of ships. Many of the newer antenna systems under development plan to use phased array antennas. The faceted nature of the AEM/S structure provides the necessary flat surfaces for mounting these future systems. The performance of the enclosed shipboard antennas is improved over conventional metallic masts because there is less blockage of the antenna. Maintenance of the enclosed antennas is reduced because the antennas are not exposed to adverse weather, wind loading, salt water, or stack gases. Less maintenance directly reduces costs over the entire service life of the ship.

AEM/S SYSTEM ATD

The AEM/S concept was demonstrated through the Office of Naval Research (ONR) AEM/S System Advanced Technology Demonstration (ATD) project. This FY 1995 ATD demonstrated the ability to design and fabricate enclosed mast structures for Navy ships. Figure 1 shows the

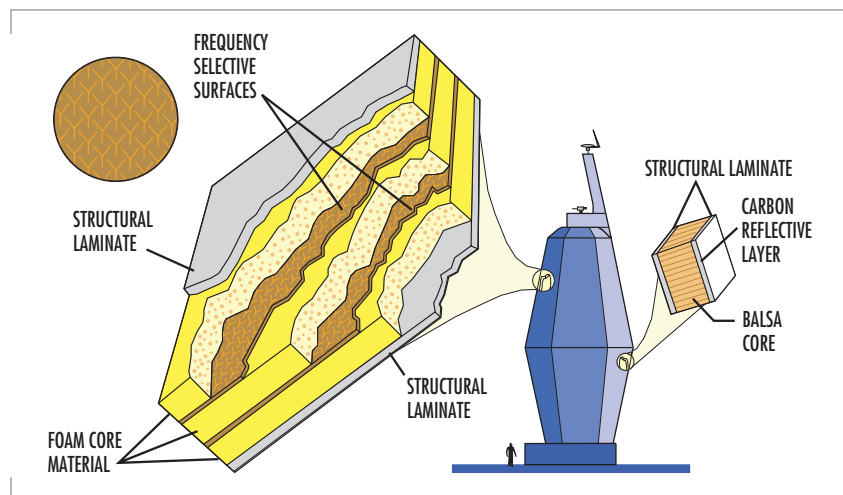


FIGURE 1. AEM/S System ATD sandwich construction concept.

AEM/S ATD configuration, with the FSS structure on the top and a balsa core reflective composite on the bottom. This ATD fused advances in electromagnetics, signature reduction, structures, materials, and manufacturing technologies. The all-composite, self-supporting enclosure is approximately 100 feet tall, 36 feet in diameter, and 40 tons in weight.

SSC San Diego played a major role in the development and the success of the AEM/S System by performing all of the electromagnetic design and development for the program. SSC San Diego's involvement included designing and validating the FSS radomes, handling antenna integration issues such as antenna placement and electromagnetic compatibility, developing new antenna designs such as the Integrated High-Frequency Antenna, and performing antenna performance predictions and measurements of major enclosed radar systems.

The FSS radome design process required artful compromise between electromagnetic, mechanical, and material engineering disciplines. Optimum mast wall design was achieved through tradeoffs between enclosed antenna system performance in the passband and the threat signal rejection level in the stop band. Also, mechanical consideration of strength bound the acceptable ranges of the composite skin and core thickness. Materials were selected for their electrical properties, mechanical strength, thermal properties, and cost.

Electromagnetic compatibility is designed into the mast through proper antenna placement. This compatibility is achieved by using the conducting decks as shielding, using the filtering characteristics of the radomes, and designing the structure to minimize electromagnetic interference while maximizing coverage.

SSC San Diego conceived the idea of mounting a high-frequency (HF) antenna to the inside surface of a radome during the research phase prior to the start of the AEM/S ATD. Eventually called the Integrated High-Frequency Antenna (IHFA), the concept offered a novel approach to the design of HF antennas for Navy ships in that (1) the radome structure provides the necessary height and volume to produce a good HF antenna, and (2) by mounting the antenna to the inside surface of the radome, the antenna cannot be seen by threat radars.

SSC San Diego also developed a new capability for antenna performance evaluation on Navy ships during the AEM/S ATD. With assistance from The Ohio State University ElectroScience Laboratory, new computer modeling tools were developed for the analysis of radome-enclosed antennas. This capability has been validated using scale-model and full-scale antenna pattern measurements.

In 1997, the AEM/S ATD culminated with the installation and at-sea testing of the mast on USS *Arthur W. Radford* (DD 968). Figure 2 shows *Arthur W. Radford* at sea with the AEM/S mast installed. This mast provides superior antenna system performance. The AEM/S mast also provides significant RCS reduction, reduced antenna system maintenance, and reduced life-cycle costs.

AEM/S FOR LPD 17

While the ATD mast was being fabricated and installed, members of the ATD project team and the Naval Sea Systems Command LPD 17 program office began discussion of potential advantages of AEM/S. Because

the ATD technology showed performance and maintenance advantages for the LPD 17 platform,* a risk mitigation project to address technology transition and design issues for LPD 17 was initiated.

As with the AEM/S System ATD, SSC San Diego has played a major role in the development and the success of the AEM/S for the LPD 17 program. SSC San Diego performed all of the electromagnetic design and development for the program. Involvement has included designing and validating the FSS radomes, handling antenna integration issues such as antenna placement and electromagnetic compatibility, developing the IHFA designs, and performing antenna performance predictions and measurements of major enclosed systems.

The design of the AEM/S for LPD 17 involves the design of two separate masts. As such, each mast has different requirements, different antenna systems, and, therefore, different challenges. In both cases, the performance requirements were more difficult to meet than those of the ATD radome, largely because of the considerable increase in the signature requirements of the masts. Another challenge for the radome for the aft mast is meeting the more stringent requirements of the SPS-48E radar. SPS-E is a higher frequency, higher gain radar that is sensitive to any variations caused by the surrounding ship structure. These requirements, in addition to the extreme structural requirements imposed by the height of the enclosure (approximately 12 meters), suggest the extraordinary interdisciplinary cooperation necessary to obtain an optimum design.

Figure 3 shows an artist's conception of the LPD 17 with the AEM/S masts installed.

Another of the many challenges associated with the LPD 17 AEM/S is the design of IHFAs for low- and high-band transmission. Under the ATD, only the design of a high-band antenna was treated. The low-band IHFA requires a radome structure that provides the necessary height and volume to produce a good HF antenna. These challenges were



FIGURE 2. AEM/S ATD at sea on USS *Arthur W. Radford* (DD 968).

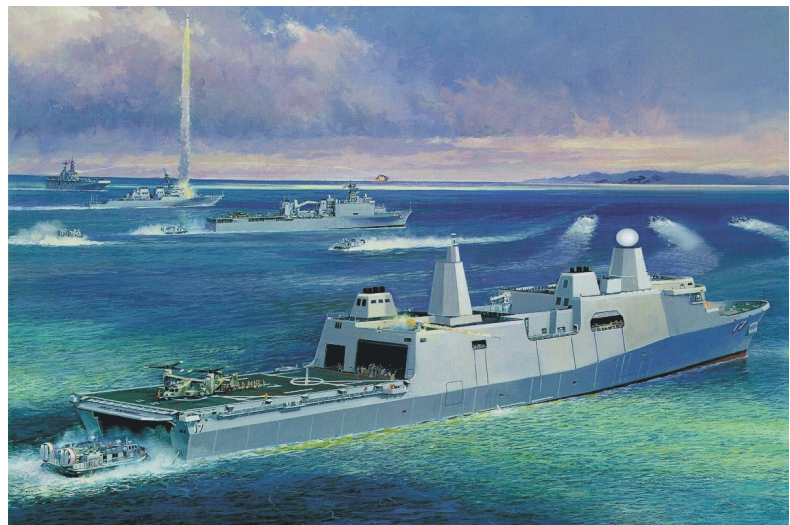


FIGURE 3. Artist's conception of the LPD 17 with AEM/S masts.

*Landing Platform Dock 17 (LPD 17), *San Antonio* class, is the latest class of amphibious force ship for the U.S. Navy. The first ship, USS *San Antonio* (LPD 17) is currently under construction.

met by the inclusion of the high-band IHFA in the shorter forward mast, and the low-band antenna in the taller aft mast.

Predicting antenna performance for enclosed antennas was particularly difficult for the SPS-48E radar. SPS-48E is a volume search radar with very high gain and low sidelobes. Predicting performance has proven to be one of the most challenging aspects of the LPD 17 design. However, a good understanding of the performance of the enclosed antenna has been obtained through advanced computer modeling and component-level measurements.

In 1999, the AEM/S risk mitigation effort culminated with the official change of the design from the contract metal masts to the AEM/S masts; this was a milestone comparable in significance to the installation of the original ATD mast on *Arthur W. Radford*. Since that time, work has continued in all areas to obtain designs that are ready to meet the production schedule of the lead ship.

CONCLUSION

The AEM/S System is a unique U.S. Navy program that has encompassed research and development, an Advanced Technology Demonstration (ATD), and new ship construction (LPD 17). This successful transition of technology has made the AEM/S System program one of the most successful programs of the last decade. The highly integrated and consensus-managed team of Navy and industry experts has made this program successful. The program's success and numerous benefits will encourage the Navy to continue implementing the AEM/S System and its associated technologies.



John H. Meloling

Ph.D. in Electrical Engineering,
Ohio State University, 1994

Current Research: Frequency-
selective surfaces; radomes;
absorbers; high-frequency
electromagnetics.

Seaweb Underwater Acoustic Nets

Joseph A. Rice, Robert K. Creber,
Christopher L. Fletcher, Paul A. Baxley,
Kenneth E. Rogers, and Donald C. Davison
SSC San Diego

INTRODUCTION

Digital signal processor (DSP) electronics and the application of digital communications theory have substantially advanced the underwater acoustic telemetry state of the art [1]. A milestone was the introduction of a DSP-based modem [2] sold as the Datasonics ATM850 [3 and 4] and later identified as the first-generation telesonar modem. To promote further development of commercial off-the-shelf (COTS) telesonar modems, the U.S. Navy invested small business innovative research (SBIR) funding and Navy laboratory support with expectations that energy-efficient, inexpensive telesonar modems would spawn autonomous undersea systems [5]. Steady progress resulted in the second-generation telesonar modem [6], marketed as the Datasonics ATM875. Encouraged by the potential demonstrated with the ATM875, the Navy funded the advanced development of a third-generation telesonar modem [7] designated the Benthos ATM885.

Seaweb is an organized network for command, control, communications, and navigation (C³N) of deployable autonomous undersea systems. Seaweb functionality implemented on telesonar hardware shows enormous promise for numerous ocean applications.

Offboard seaweb nodes of various types may be readily deployed from high-value platforms including submarine, ship, and aircraft, or from unmanned undersea vehicles (UUVs) and unmanned aerial vehicles (UAVs). The architectural flexibility afforded by seaweb wireless connections permits the mission planner to allocate an arbitrary mix of node types with a node density and area coverage appropriate for the given telesonar propagation conditions and for the mission at hand.

The initial motivation for seaweb is a requirement for wide-area undersea surveillance in littoral waters by means of a deployable autonomous distributed system (DADS) such as that shown in Figure 1. Future sensor nodes in a DADS network generate concise antisubmarine warfare (ASW) contact reports that seaweb will route to a master node for field-level data fusion [8]. The master node communicates with manned command centers via gateway nodes such as a sea-surface buoy radio-linked with space satellite networks, or a ship's sonar interfaced to an onboard seaweb server.

DADS operates in 50- to 300-m waters with node spacing of 2 to 5 km. Primary network packets are contact reports with about 1000 information

ABSTRACT

Seaweb networks use digital signal processor (DSP)-based telesonar underwater acoustic modems to interconnect fixed and mobile nodes. Backbone nodes are autonomous, stationary sensors and telesonar repeaters. Peripheral nodes include unmanned undersea vehicles (UUVs) and specialized devices such as low-frequency sonar projectors. Gateway nodes provide interfaces with command centers afloat, submerged, ashore, and aloft, including access to terrestrial, airborne, and space-based networks. Seaweb command, control, communications, and navigation (C³N) technology coordinates deployable assets for accomplishing given missions in littoral ocean environments. A series of annual experiments drives seaweb technology development by implementing increasingly sophisticated wide-area networks of deployable autonomous undersea sensors.

bits [9]. DADS sensor nodes asynchronously produce these packets at a variable rate dependent on the receiver operating characteristics for a particular sensor suite and mission.

Following *ad hoc* deployments, DADS relies on the seaweb network for self-organization including node identification, clock synchronization on the order of 0.1 to 1.0 s, node geolocalization on the order of 100 m, assimilation of new nodes, and self-healing following node failures. Desired network endurance is up to 90 days.

DADS is a fixed grid of inexpensive interoperable nodes. This underlying cellular network architecture is well suited for supporting an autonomous oceanographic sampling network (AOSN) [10], including C³N for autonomous operations with UUV mobile nodes.

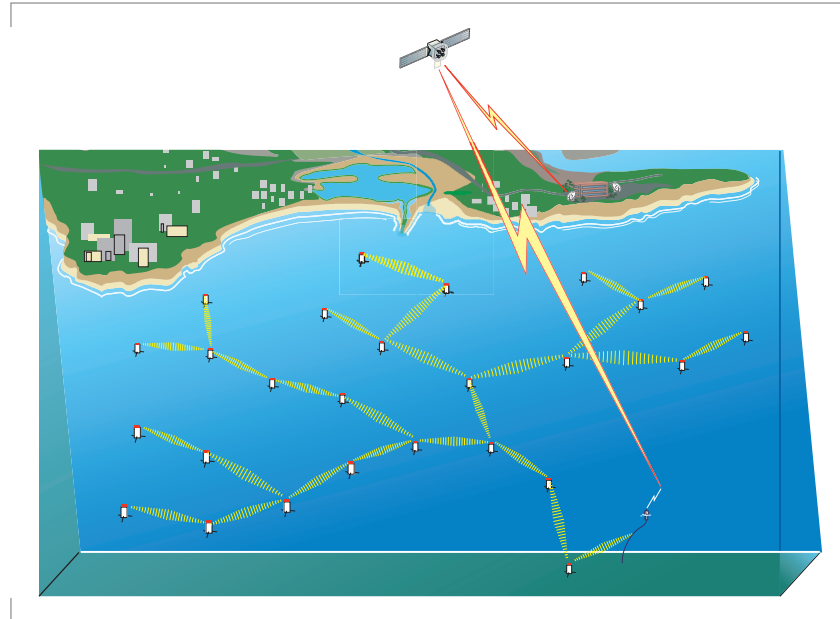


FIGURE 1. Seaweb underwater acoustic networking enables C³N for DADS and other deployable autonomous undersea systems. Gateways to manned control centers include radio links to space or shore and telesonar links to ships.

CONCEPT OF OPERATIONS

Telesonar wireless acoustic links interconnect distributed undersea instruments, potentially integrating them as a unified resource and extending "net-centric" operations into the undersea environment.

Seaweb is the realization of such an undersea wireless network [11] of fixed and mobile nodes, including various interfaces to manned command centers. It provides the C³N infrastructure for coordinating appropriate assets to accomplish a given mission in an arbitrary ocean environment.

The seaweb backbone is a set of autonomous, stationary nodes (e.g., deployable surveillance sensors, repeaters). Seaweb peripherals are mobile nodes (e.g., UUVs, including swimmers, gliders, and crawlers) and specialized nodes (e.g., bistatic sonar projectors).

Seaweb gateways provide connections to command centers submerged, afloat, aloft, and ashore. Telesonar-equipped gateway nodes interface seaweb to terrestrial, airborne, and space-based networks. For example, a telesonobuoy serves as a radio/acoustic communications (racom) interface, permitting satellites and maritime patrol aircraft to access submerged, autonomous systems. Similarly, submarines can access seaweb with telesonar signaling through the WQC-2 underwater telephone band or other high-frequency sonars [12]. Seaweb provides the submarine commander with digital connectivity at speed and depth and with bidirectional access to all seaweb-linked resources and distant gateways.

A seaweb server resides at manned command centers and is the graphical user interface to the undersea network as shown in Figure 2. The server

archives all incoming data packets and provides read-only access to client stations via the Internet. A single designated "super" server controls and reconfigures the network.

Low-bandwidth, half-duplex, high-latency telesonar links limit seaweb quality of service.

Occasional outages from poor propagation or elevated noise levels can disrupt telesonar links [13]. Ultimately, the available energy supply dictates service life, and battery-limited nodes must be energy conserving [14]. Moreover, seaweb must ensure transmission security by operating with low bit-energy per noise-spectral-density (E_b/N_0) and by otherwise limiting interception by unauthorized receivers. Seaweb must therefore be a revolutionary information system bound by these constraints.

Simplicity, efficiency, reliability, and security are the governing design principles. Half-duplex handshaking [15] asynchronously establishes adaptive telesonar links [16] as shown in Figure 3. The initiating node transmits a request-to-send (RTS) waveform with a frequency-hopped, spread-spectrum (FHSS) [17] pattern or direct-sequence spread-spectrum (DSSS) [18] pseudo-random carrier uniquely addressing the intended receiver. (Alternatively, the initiating node may transmit a universal code for broadcasting or when establishing links with unknown nodes.) The addressed node detects the request and awakens from an energy-conserving sleep state to demodulate. Further processing of the RTS signal provides an estimate of the channel scattering function and signal excess. The addressed node then acknowledges receipt with a FHSS or DSSS acoustic response. This clear-to-send (CTS) reply specifies appropriate modulation parameters for the ensuing message packets based upon the measured channel conditions. Following this RTS/CTS handshake, the initiating node transmits the data packet(s) with nearly optimal bit-rate, modulation, coding, and source level.

Spread-spectrum modulation is consistent with the desire for asynchronous multiple access to the physical channel using code-division multiple-access (CDMA) networking [19]. Nevertheless, the seaweb concept

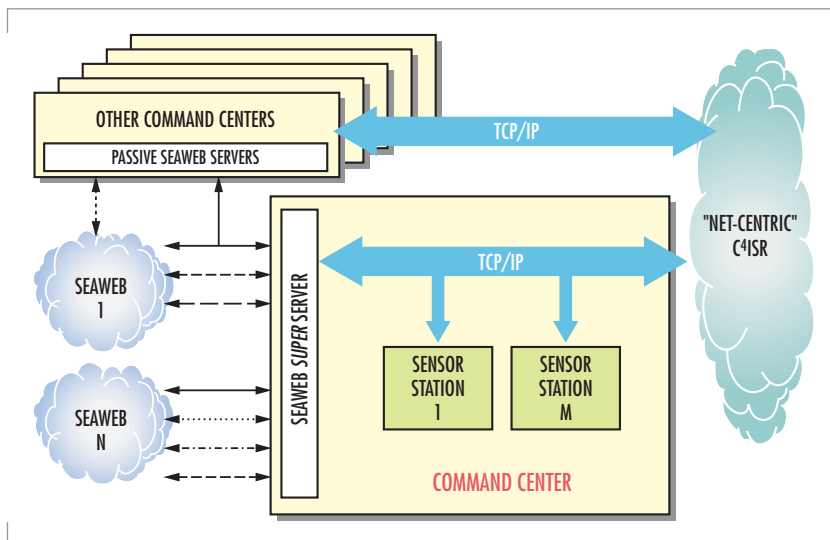


FIGURE 2. Seaweb extends modern "net-centric" interconnectivity to the undersea realm. Wireless underwater networks include gateway nodes with radio, acoustic, wire, or fiber links to manned command centers where a seaweb server provides a graphical user interface. Command centers may be aboard ship, submarine, aircraft, or ashore. They may be geographically distant and connected to the gateway node via space satellite or terrestrial Internet. At the designated command center, a seaweb "super" server manages and controls the undersea network. All seaweb servers archive seaweb packets and provide data access to client stations via the Internet. A single designated super server controls and reconfigures the network.

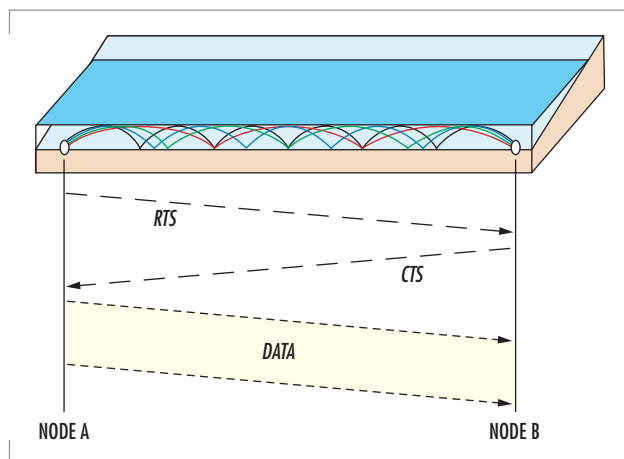


FIGURE 3. Telesonar handshake protocol for data transfer involves Node A initiating a request-to-send (RTS) modulated with a channel-tolerant, spread-spectrum pattern uniquely associated with intended receiver, Node B. So addressed, Node B awakens and demodulates the fixed-length RTS packet. Node B estimates the channel parameters using the RTS as a probe signal. Node B responds to A with a fixed-length clear-to-send (CTS) that fully specifies the modulation parameters for the data transfer. Node A then sends the data packet(s) with optimal source level, bit-rate, modulation, and coding. If Node B receives corrupted data, it initiates a selective automatic repeat request (ARQ).

does not exclude time-division multiple-access (TDMA) or frequency-division multiple-access (FDMA) methods, and is pursuing hybrid schemes suited to the physical-layer constraints. In a data transfer, for example, the RTS/CTS exchange might occur as an asynchronous CDMA dialog in which the data packets are queued for transmission during a time slot or within a frequency band such that collisions are avoided altogether.

The seaweb architecture of interest includes the physical layer, the media-access-control (MAC) layer, and the network layer. These most fundamental layers of communication functionality support higher layers that will tend to be application-specific.

At the physical layer, an understanding of the transmission channel is obtained through at-sea measurements and numerical propagation models. Knowledge of the fundamental constraints on telesonar signaling translates into increasingly sophisticated modems. DSP-based modulators and demodulators permit the application of modern digital communication techniques to exploit the unique aspects of the underwater channel. Directional transducers further enhance the performance of these devices [20].

The MAC layer supports secure, low-power, point-to-point connectivity, and the telesonar handshake protocol is uniquely suited to wireless half-duplex networking with slowly propagating channels. Handshaking permits addressing, ranging, channel estimation, adaptive modulation, and power control. The seaweb philosophy mandates that telesonar links be environmentally adaptive [21], with provision for bidirectional asymmetry.

Network supervisory algorithms can execute either at an autonomous master node or at the seaweb server. Seaweb provides for graceful failure of network nodes, addition of new nodes, and assimilation of mobile nodes. Essential by-products of the telesonar link are range measurement, range-rate measurement, and clock-synchronization. Collectively, these C³N features support network initialization, node localization, route configuration, resource optimization, and maintenance.

DEVELOPMENTAL APPROACH

Given the DADS performance requirements, seaweb research is advancing telesonar modem technology for reliable underwater signaling by addressing the issues of (1) adverse transmission channel; (2) asynchronous networking; (3) battery-energy efficiency; (4) transmission security; and (5) cost.

Despite an architectural philosophy emphasizing simplicity, seaweb is a complex system and its development is a grand challenge. The high cost of sea testing and the need for many prototype nodes motivate extensive engineering system analysis following the ideas of the previous section.

Simulations using an optimized network engineering tool (OPNET) with simplified ocean acoustic propagation assumptions permit laboratory refinement of networking protocols [22] and initialization methods [23]. Meanwhile, controlled experimentation in actual ocean conditions incrementally advances telesonar signaling technology [24].

Seaweb experiments implement the results from these research activities with a periodic concentration of resources in prolonged ocean experiments.

The annual seaweb experiments validate system analysis and purposefully evolve critical technology areas such that the state-of-the-art advances with greater reliability, functionality, and quality of service. The objective of the seaweb experiments is to exercise telesear modems in networked configurations where various modulation and networking algorithms can be assessed. In the long-term, the goal is to provide for a self-configuring network of distributed assets, with network links automatically adapting to the prevailing environment through selection of the optimum transmit parameters.

A full year of hardware improvements and in-air network testing helps ensure that the incremental developments tested at sea will provide tractable progress and mitigate overall developmental risk. In particular, DADS relies on the annual seaweb engineering experiments to push telesear technology for undersea wireless networking. After the annual seaweb experiment yields a stable level of functionality, the firmware product can be further exercised, and refinements can be instituted during DADS system testing and by spin-off applications throughout the year. For example, in year 2001, seaweb technology enables the March–June *FRONT-3* ocean observatory on the continental shelf east of Long Island, NY [25]. These applications afford valuable long-term performance data that are not obtainable during seaweb experiments when algorithms are in flux and deployed modems are receiving frequent firmware upgrades.

The *Seaweb* '98, '99, and 2000 operating area in Buzzards Bay is framed in Figure 4. An expanse of 5- to 15-m shallow water is available for large-area network coverage with convenient line-of-sight radio access to Datasonics and Benthos facilities in western Cape Cod, MA. A shipping channel extending from the Bourne Canal provides episodes of high shipping noise useful for stressing the link signal-to-noise ratio (SNR) margins. The seafloor is patchy with regions of sand, gravel, boulders, and exposed granite.

Figure 5 shows *Seaweb* '98, '99, and 2000 modem rigging. Experiments occur during August and September when weather is conducive to regular servicing of deployed network nodes.

A representative sound-speed profile inferred from a conductivity-temperature-depth (CTD) probe during *Seaweb* '98 is shown in Figure 6. For observed August and September sound-speed profiles, ray tracing suggests maximum direct-path propagation to ranges less than 1000 m, as Figure 7 shows. Beyond this distance, received acoustic energy is via boundary forward scattering. Ray tracing further indicates that received signal energy at significant ranges is attributable



FIGURE 4. The test site for *Seaweb* '98, '99, and 2000 is northern Buzzards Bay, MA. Water depth is 5 to 15 m.

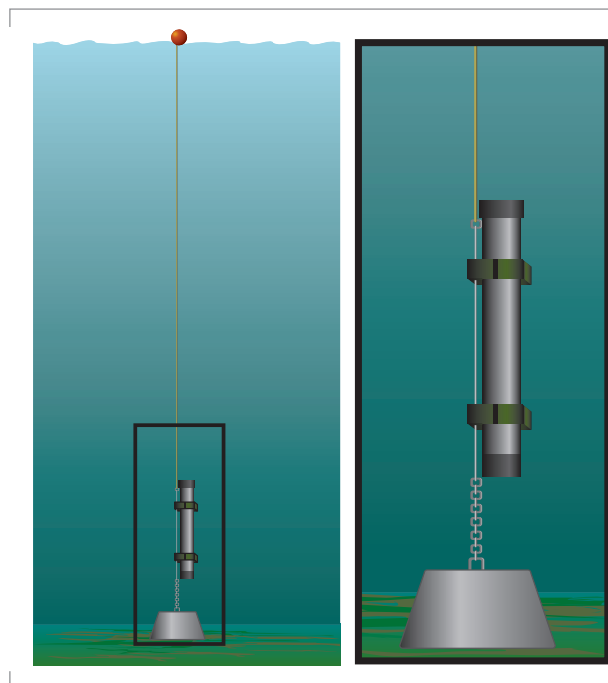


FIGURE 5. *Seaweb* '98, '99, and 2000 modems are deployed in Buzzards Bay with concrete weight, riser line, and surface float. The shallow water and simple rigging permit a small craft to rapidly service the network. Servicing includes battery replacement and modem firmware downloads.

to a very small near-horizontal continuum of projector elevation launch angles. Figure 8 presents predicted impulse responses for 10 ranges, each revealing multipath spreads of about 10 ms [26]. All ranges are considered "long" with respect to water depth. Summer afternoon winds and boat traffic regularly roughen the sea surface, increasing scattering loss and elevating noise levels.

SEAWEB '98 EXPERIMENT

Seaweb '98 led off a series of annual ocean experiments intended to progressively advance the state of the art for asynchronous, non-centralized networking. *Seaweb '98* used the Datasonics ATM875 second-generation telesonar modem [27] recently available as the product of a Navy SBIR Phase-2 contract.

The ATM875 normally uses 5 kHz of acoustic bandwidth with 120 discrete multiple-frequency shift keying (MFSK) bins configured to carry six Hadamard codewords of 20 tones each. Interleaving the codewords across the band increases immunity to frequency-selective fading, and Hadamard coding yields a frequency diversity factor of 5 for adverse channels having low or modest spectral coherence. This standard ATM875 modulation naturally supports three interleaved FDMA sets of 40 MFSK tonals and two codewords each. To further reduce multi-access interference (MAI) between sets, half the available bandwidth capacity provided additional guardbands during *Seaweb '98*. Thus, only 20 MFSK tonals composing one Hadamard codeword formed each FDMA set. The *Seaweb '98* installation was three geographic clusters of nodes with FDMA sets "A" through "C" mapped by cluster. For example, all nodes in cluster A were assigned the same FDMA carrier set for reception. Each cluster contained a commercial oceanographic sensor at a leaf node asynchronously introducing data packets into the network. This FDMA architecture was an effective multi-access strategy permitting simultaneous network activity in all three clusters without MAI [28]. A drawback of FDMA signaling is the inefficient use of available bandwidth. *Seaweb '98* testing was based on a conservative 300-bit/s modulation to yield a net FDMA bit-rate of just 50 bit/s. This was an acceptable rate since the *Seaweb '98* objective was to explore networking concepts without excessive attention to signaling issues. Within a cluster, TDMA was the general rule broken only by deliberate intrusion from the command center.

The gateway node is an experimental Navy racom buoy (Figure 9). The "master" node was installed approximately 1500 m from the gateway node. Gateway and master nodes formed cluster C, and so received and demodulated only the FDMA carriers of set C. Exercising the link between gateway and master nodes during various multi-hour and multi-day

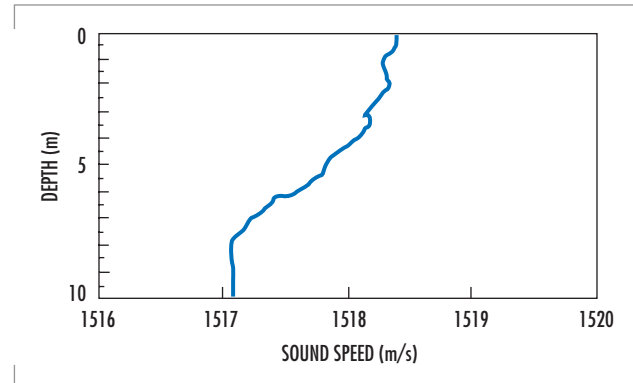


FIGURE 6. Sound-speed profiles calculated from conductivity and temperature probes are generally downward refracting during August–September at the *Seaweb '98*, '99, and 2000 site. This sound-speed profile, 1 of 14 obtained during *Seaweb '98*, is typical. The sound-speed gradient evident here is caused by summertime sea-surface warming.

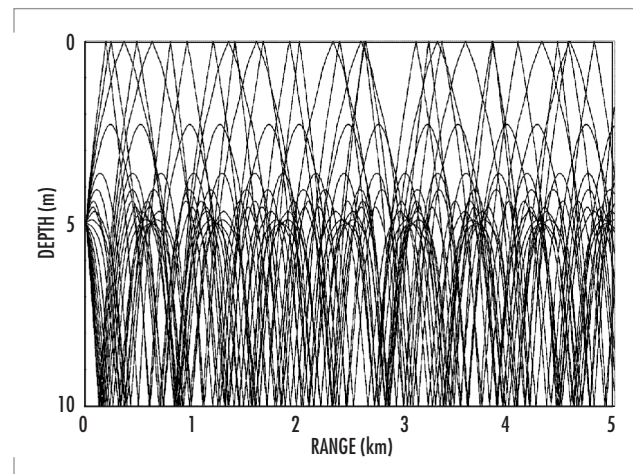


FIGURE 7. *Seaweb '98* propagation refracts downward in response to the vertical sound-speed gradient observed in Figure 6. Rays traced from a $\pm 2.5^\circ$ vertical fan of launch elevation angles model the telesonar sound channel for transmitter at 5-m depth. A parametric modeling study assessing the dependence of modem depth for this environment confirmed the general rule that long-range signaling in downward-refracting, non-ducted waters is optimized with modems placed nearer the seafloor. Hence, *Seaweb '98*, '99, and 2000 modem transducers are generally about 2 m above the bottom.

periods yielded link statistics for improving the wake-up and synchronization schemes in the modem receiver acquisition stage. This point-to-point testing identified specific suspected problems in the fledgling ATM875 implementation, and firmware modifications improved the success of packet acquisition from 80% to 97%.

Installation of a three-node subset of cluster A added a relay branch around Scraggy Neck, a peninsula protruding into Buzzards Bay. An Ocean Sensors CTD produced data packets relayed via each of the intervening A nodes to the master node, and then on to the gateway node. Each relay link was about 1500 m in range. Direct addressing of cluster-A nodes from the gateway node confirmed the existence of reliable links to all but the outermost node. Remarkably, a reliable link existed between two nodes separated by 3.6 km in spite of shoaling to 1 to 2 m in intervening waters! Various network interference situations were intentionally and unintentionally staged and tested until this simple but unprecedented relay geometry was well understood.

These early tests realized an unexpected benefit of the gateway link between the racom buoy and the radio-equipped workboat. End-to-end functionality of a newly installed node could be immediately verified. Field personnel would use a deck unit and the gateway node to test the network circuit that included the new modem as an intermediate node, or they would bidirectionally address the new modem via just the gateway route. Effectively, the workboat was a mobile node in the network equipped with both telesonar and gateway connections.

At this point, associates from the National Oceanic and Atmospheric Agency (NOAA) and Naval Surface Warfare Center (NSWC) visited *Seaweb* '98. A boat delivered them far into Buzzards Bay, where a hydrophone (deployed over the side with a telesonar deck unit) turned

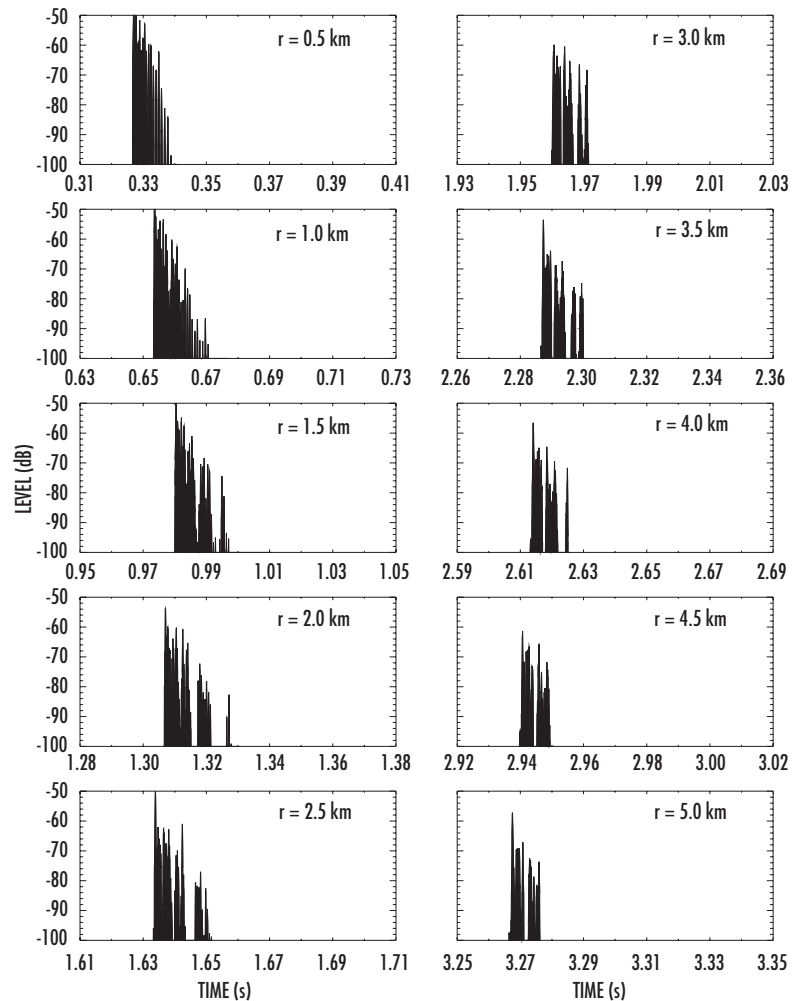


FIGURE 8. For a 10-m deep *Seaweb* '99 channel, a 2-D Gaussian beam model predicts impulse responses for receivers located at 10 ranges, r . Response levels are in decibels referenced to a 0-dB source. Multipath spread is about 10 ms. Note the *Seaweb* '98, '99, and 2000 working ranges are hundreds of times greater than the water depths, and boundary interactions are therefore complex. For rough sea floor and sea surface, the 2-D model approximation must give way to 3-D forward scattering, and the predicted response structures will instead be smeared by out-of-plane propagation. *Seaweb* 2000 testing includes channel probes designed to directly measure channel scattering functions with receptions recorded at various ranges by telesonar test beds. These channel measurements support analysis of experimental signaling and help calibrate an experimental 3-D Gaussian beam model under development for telesonar shallow-water performance prediction.

the boat into just such a mobile network node. The visitors typed messages, which were relayed through the network and answered by personnel at the ashore command center.

Next, a branch was added to cluster A with a Falmouth Scientific 3-D current meter and CTD. Network contention was studied by having the two cluster-A sensor nodes generate packets at different periods such that network collisions would occur at regular intervals with intervening periods of non-colliding network activity.

Finally, cluster B was introduced to the network with internode separations of 2 km. A third device generated data packets. With all available network nodes installed and functioning, the remaining few days involved a combination of gradually arranging network nodes with greater spacing as charted in Figure 10, and of doing specialized signal testing with the telesonar test bed [29]. In addition, the telesonar test bed was deployed in the center of the network for five data-acquisition missions and recorded 26 hours of acoustic network activity. The test bed also included a modem, permitting it to act as the tenth network node and giving ashore operators the ability to remotely control and monitor test-bed operations. The test-bed node provides raw acoustic data for correlation with automatic modem diagnostics, providing opportunity to study failure modes using recorded time series.

Seaweb '98 demonstrated the feasibility of low-cost distributed networks for wide-area coverage. During 3 weeks of testing in September, the network performed reliably through a variety of weather and noise events. Individual network links spanned horizontal ranges hundreds of water depths in length. The *Seaweb '98* network connected widely spaced autonomous modems in a binary-tree topology with a master node at the base and various oceanographic instruments at outlying leaf nodes. Also connected to the master node was an acoustic link to a gateway buoy, providing a line-of-sight digital radio link to the command center ashore. The network transported data packets acquired by the oceanographic instruments through the network to the master node, on to the gateway node, and then to the command center. The oceanographic instruments and modems generally operated according to preprogrammed schedules designed to periodically produce network collisions, and personnel at the command center or aboard ship also remotely controlled network nodes in an asynchronous manner.

The most significant result of *Seaweb '98* is the consistent high quality of received data obtained from



FIGURE 9. In *Seawebs '98, '99, and 2000*, a radio/acoustic communications (racom) buoy provides a very reliable line-of-sight packet-radio link to seaweb servers at the shore command center and on the work boat. The radio link is a 900-MHz spread-spectrum technology commercially known as Freewave. In *Seawebs '99 and 2000*, additional gateway nodes using cellular modems linked via Bell Atlantic and the Internet provide even greater flexibility and provide access by seaweb servers at various locales across the country.

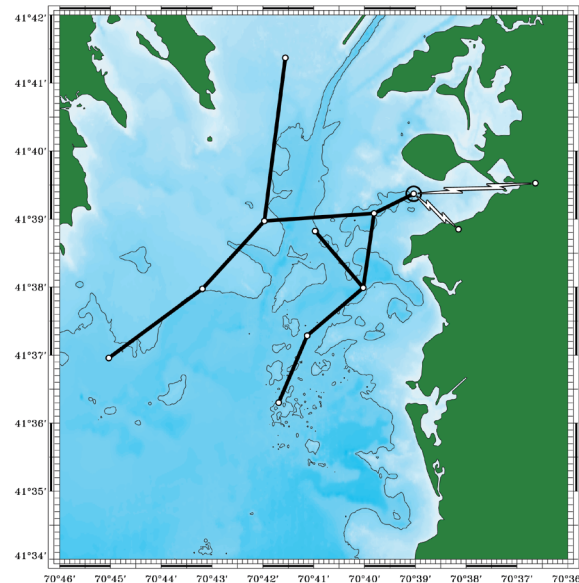


FIGURE 10. *Seaweb '98* demonstrated store and forward of data packets from remote commercial sensors including a CTD, a current vector meter, and a tilt/heaving sensor (at the most northerly, westerly, and southerly leaf nodes, respectively) via multiple network links to the racom gateway buoy (large circle). Data packets are then transmitted to the ashore command center via line-of-sight packet radio. An FDMA network with three frequency sets reduced the possibility of packet collisions. Following extensive firmware developments supported by this field testing, the depicted topology was exercised during the final days of the experiment. Isobaths are contoured at 5-m intervals.

remote autonomous sensors. Data packets arrived at the command center via up to four acoustic relays and one RF relay. About 2% of the packets contained major bit-errors attributable to intentional collisions at the master node. The quality of data was very high even after the network was geographically expanded. Reliable direct telesonar communications from the gateway node to a node nearly 7 km distant suggested the network could be expanded considerably more, in spite of the non-ducted 10-m deep channel. The *Seaweb* '98 environment could have supported 4-km links using the same ATM875 modems and omnidirectional transducers. Attesting to the channel-tolerant nature of the MFSK modulation, an early phase of testing maintained a 3-km link between two nodes separated by a 1- to 2-m deep rocky shoal. Consistent network degradation occurred during most afternoons and is attributable to summer winds roughening the sea-surface boundary and thus scattering incident acoustic energy. Automated network operations continued during heavy rains and during ship transits through the field.

Seaweb '98 demonstrated the following network concepts: (1) store and forward of data packets; (2) transmit retries and automatic repeat request; (3) packet routing; and (4) cell-like FDMA node grouping to minimize MAI between cells. In addition, the following DADS concepts were demonstrated: (1) networked sensors; (2) wide-area coverage; (3) racom gateway; (4) robustness to shallow-water environment; (5) robustness to shipping noise; (6) low-power node operation with sleep modes; (7) affordability; and (8) remote control. Finally, *Seaweb* '98 resulted in dramatic improvements to the ATM875 modem and improved its commercial viability for non-networked applications.

Seaweb '98 observations underscore the differences between acoustic networks and conventional networks. Limited power, low bandwidth, and long propagation times dictate that seaweb networks be simple and efficient. Data compression, forward error correction, and data filtering must be employed at the higher network levels to minimize packet sizes and retransmissions. At the network layer, careful selection of routing is required to minimize transmit energy, latency, and net energy consumption, and to maximize reliability and security. At the physical and MAC layers, adaptive modulation and power control are the keys to maximizing both channel capacity (bit/s) and channel efficiency (bit-km/joule).

SEAWEB '99 EXPERIMENT

Seaweb '99 continued the annual series of telesonar experiments incrementally advancing the state of the art for undersea wireless networks. During a 6-week period, up to 15 telesonar nodes operated in various network configurations in the 5- to 15-m waters of Buzzards Bay. Network topologies provided compound multi-link routes. All links used a rudimentary form of the telesonar handshake protocol featuring an adaptive power-control technique for achieving sufficient but not excessive SNR at the receiver. Handshaking provided the means for resolving packet collisions automatically using retries from the transmitter or automatic-repeat-request (ARQ) packets from the receiver.

The multi-access strategy was a variation of FDMA wherein the six available 20-tone Hadamard words provided six separate FDMA sets, A through F. Rather than clustering the FDMA sets as in *Seaweb* '98, the notion here was to optimally assign FDMA receiver frequencies to the various nodes in an attempt to minimize collisions through spatial

separation and the corresponding transmission loss. This approach represents an important step toward network self-configuration and prefigures the future incorporation of secure CDMA spread-spectrum codes to be uniquely assigned to member nodes during the initialization process.

Node-to-node ranging employed a new implementation of a round-trip-travel time measurement algorithm with 0.1-ms resolution linked to the DSP clock rate. Range estimation simply assuming a constant 1500 m/s sound speed was consistently within 5% of GPS-based measurements for all distances and node pairs.

A significant development was the introduction of the seaweb server. It interprets, formats, and routes downlink traffic destined for undersea nodes. On the uplink, it archives information produced by the network, retrieves the information for an operator, and provides database access for client users. The server manages seaweb gateways and member nodes. It monitors, displays, and logs the network status. The server manages the network routing tables and neighbor tables and ensures network interoperability. *Seaweb '99* modem firmware permitted the server to remotely reconfigure routing topologies, a foreshadowing of future self-configuration and dynamic network control. The seaweb server is a graphical set of LabView virtual instruments implemented under Windows NT on a laptop PC. A need for the server was illustrated when operators bypassed server oversight and inadvertently produced a circular routing where a trio of nodes continuously passed a packet between themselves until battery depletion finally silenced the infinite loop.

In *Seaweb '99*, the server simultaneously linked with a Bell Atlantic cellular digital packet data (CDPD) gateway node via the Internet and with the packet-radio racom gateway link via a serial port. A milestone was the establishment of a gateway-to-gateway route through the seaweb server that was exercised automatically over a weekend.

Seaweb '99 included an engineering test for the "Front-Resolving Observation Network with Telemetry" (FRONT) application, with large acoustic Doppler current profiler (ADCP) data packets synthesized and passed through the network with TDMA scheduling. A study of network capacity examined the periodic uplinking of data packets while asynchronously issuing server-generated downlink commands to poll sensors.

For every packet received by a *Seaweb '99* node, the modem appended link metrics such as bit-error rate (BER), automatic gain control (AGC), and SNR. These diagnostics aided post-mortem system analysis. Performance correlated strongly with environmental factors such as refraction, bathymetry, wind, and shipping, although no attempt was made to quantify these relationships in *Seaweb '99*.

The ATM875 second-generation telesonar modem again served as the workhorse modem for all network nodes. During the last phase of the experiment, progress was thwarted by memory limitations of the Texas Instruments TMS320C50 DSP. A firmware bug could not be adequately resolved because of lack of available code space for temporary in-line diagnostics. Consequently, the final days of the test reverted to a prior stable version of the *Seaweb '99* code and the 15-node network charted in Figure 11 covered a less ambitious area than originally intended. These limitations plus the desire to begin implementing FHSS and DSSS signaling motivated the initiation of ATM885 third-generation telesonar modem development for *Seaweb 2000*.

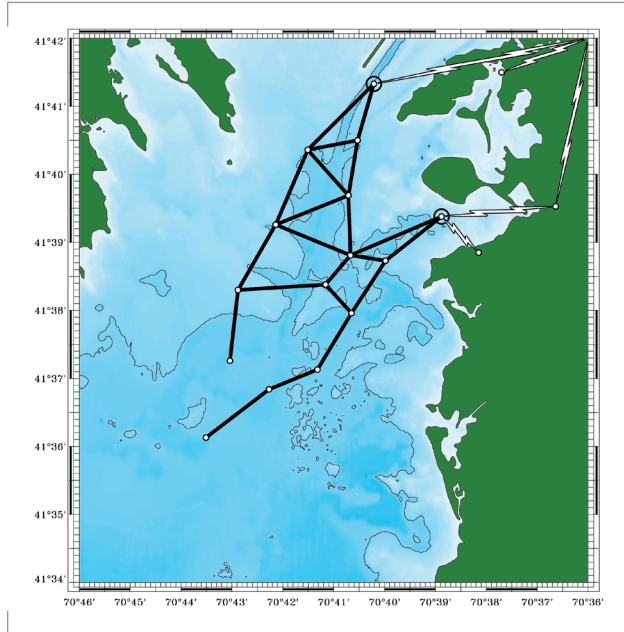


FIGURE 11. *Seaweb '99* explored the use of handshaking and power control. An ADCP sensor node, a tilt/heaving sensor node, and a CTD sensor node generated data packets, and the network routed them through various paths. The racom gateway node (easterly large circle) again provided a solid link to shore. A second gateway node (northerly large circle) installed on a Coast Guard caisson near the Bourne canal provided a Bell Atlantic cellular modem link to the Internet and then to the command center. The Seaweb server running on a laptop PC managed both gateway connections and archived all network activity.

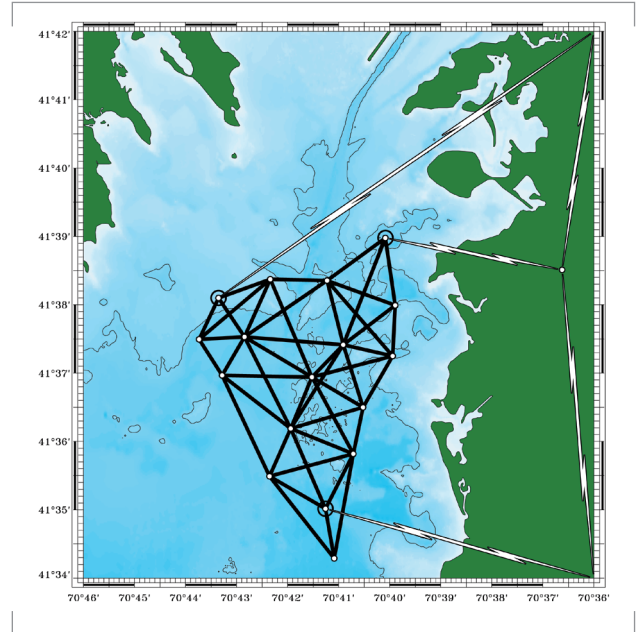


FIGURE 12. *Seaweb 2000* exercised the telesonar handshake protocol in a network context. The 17-node seaweb network delivered oceanographic data from sensor nodes to gateway nodes—one with line-of-sight packet radio and two with cellular telephone modems. During the final week, a seaweb super server operating at the Oceans 2000 Conference in Providence, RI, administered *Seaweb 2000* via the Internet.

SEAWEB 2000 EXPERIMENT

The *Seaweb 2000* network included up to 17 nodes, with one of the fully connected configurations charted in Figure 12. This experiment achieved major advances in both hardware and firmware.

Use of the ATM875 modem during *Seawebs '98* and *'99* continually thwarted progress in firmware development because of limited memory and processing speed. The ATM885 modem shown in Figure 13 overcomes these shortcomings with the incorporation of a more powerful DSP and additional memory. Now, telesonar firmware formerly encoded by necessity as efficient machine language is reprogrammed on the ATM885 as a more structured set of algorithms. The *ForeFRONT-1* (November 1999), *FRONT-1* (December 1999), *ForeFRONT-2* (April 2000), *Sublink 2000* (May 2000), and *FRONT-2* (June 2000) experiments hastened the successful transition of *Seaweb '99* firmware from the ATM875 to the ATM885. These intervening seaweb applications were stepping stones toward achieving basic ATM885 hardware readiness prior to instituting *Seaweb 2000* upgrades.

Seaweb 2000 implements in firmware the core features of a compact, structured protocol. The protocol efficiently maps network-layer and MAC-layer functionality onto a physical layer based on channel-tolerant, 64-bit utility packets and channel-adaptive, arbitrary-length data packets. Seven utility packet types are implemented for *Seaweb 2000*. These packet types permit data transfers and node-to-node ranging. A richer set

of available utility packets is being investigated with OPNET simulations, but the seven core utility packets provide substantial networking capability.

The initial handshake consists of the transmitter sending an RTS packet and the receiver replying with a CTS packet. This roundtrip establishes the communications link and probes the channel to gauge optimal transmit power. Future enhancements to the protocol will support a choice of data modulation methods, with selection based on channel estimates derived from the RTS role as a probe signal. A "busy" packet is issued in response to an RTS when the receiver node decides to defer data reception in favor of other traffic. Following a successful RTS/CTS handshake, the data packet(s) are sent. The *Seaweb 2000* core protocol provides for acknowledgments, either positive or negative, of a data message. The choice of acknowledgment type will depend on the traffic patterns associated with a particular network mission. *Seaweb 2000* explored the factors that will guide this application-specific choice.

A "ping" utility packet initiates node-to-node and node-to-multinode identification and ranging. An "echo" packet is the usual response to a received ping.

In *Seaweb 2000*, FDMA architectures are superseded by hybrid CDMA/TDMA methods for avoiding mutual interference. FDMA methods sacrifice precious bandwidth and prolong the duration of a transmission, often aggravating MAI rather than resisting it. Furthermore, the use of a small number of frequency sets is viewed as an overly restrictive networking solution. Although these drawbacks were expected, *Seawebs '98* and *'99* employed FDMA primarily for ease of implementation as a simple extension to the rigid ATM875 telesonar machine code. The ATM885 permits a break from those restrictions.

Seaweb 2000 execution fully incorporates the experimental approach tried in *Seaweb '99* of establishing two parallel networks—one in air at the command center and one in the waters of Buzzards Bay. This approach minimizes time-consuming field upgrades by providing a convenient network for troubleshooting deployed firmware and testing code changes prior to at-sea downloads.

As a further analysis aid, all modems now include a data-logging feature. All output generated by the ATM885 and normally available via direct serial connection is logged to an internal buffer. Thus, the behavior of autonomous nodes can be studied in great detail after recovery from the sea. To take maximum advantage of this capability, *Seaweb 2000* code includes additional diagnostics related to channel estimation (e.g., SNR, multipath spread, Doppler spread, range rate, etc.), demodulation statistics (e.g., BER, AGC, intermediate decoding results, power level, etc.), and networking (e.g., data packet source, data packet sink, routing path, etc.). For seaweb applications, the data-logging feature can also support the archiving of data until such time that an adjacent node is able to download the data. For example, a designated sink node operating without access to a gateway node can collect all packets forwarded from the

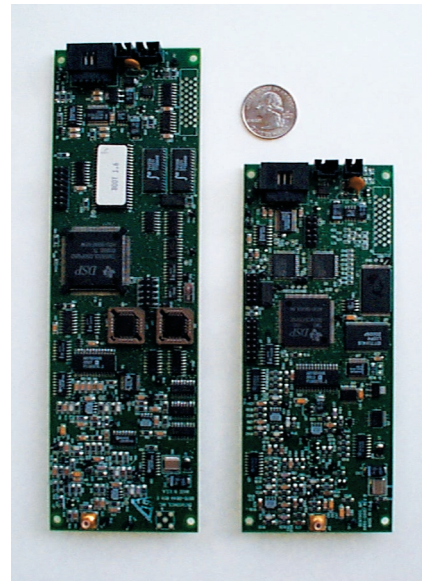


FIGURE 13. The TMS320C5410-based ATM885 telesonar modem debuted in *Seaweb 2000* with a four-fold increase in memory and processing speed over the TMS320C50-based ATM875. This hardware upgrade reduced battery-energy consumption and overcame firmware-development limitations experienced in *Seaweb '99*. The ATM885 supports 100 million instructions per second (MIPS) and 320K words of memory compared with 25 MIPS and 74K available from the ATM875.

network and telemeter them to a command center when interrogated by a gateway (such as a ship arriving on station for just such a data download).

Increasing the value of diagnostic data, the C5410 real-time clock is maintained even during sleep state. Although this clock may not have the stability required for certain future network applications, its availability permits initial development of in-water clock-synchronization techniques.

The new ATM885 modem also includes a provision for a watchdog function hosted aboard a microchip independent of the C5410 DSP. The watchdog resets the C5410 DSP upon detection of supply voltage drops or upon cessation of DSP activity pulses. The watchdog provides a high level of fault tolerance and permits experimental modems to continue functioning in spite of system errors. A watchdog reset triggers the logging of additional diagnostics for thorough troubleshooting after modem recovery.

An aggressive development schedule following *Seaweb '99* and preceding *Seaweb 2000* matured the seaweb server as a graphical user interface with improved reliability and functionality consistent with *Seaweb 2000* upgrades.

Recent telesonar engineering tests have played host to an applied research effort known as SignalEx [30]. This research uses the telesonar test beds to record high-fidelity acoustic receptions and measure relative performance for numerous signaling methods. *Seaweb 2000* hosted SignalEx testing during the second week of testing. The advantage of coupling SignalEx research with seaweb engineering is that both activities benefit—SignalEx gains resources and seaweb gains added empirical test control. By the fifth week, the major *Seaweb 2000* engineering developments reached a level of stability permitting experimental use of acoustic navigation methods for node localization, cost functions for optimized network routing, and statistics gathering for network traffic analysis.

In summary, the specific implementation objectives of *Seaweb 2000* are (1) packet forwarding through network, under control of remotely configurable routing table; (2) 64-bit header; (3) improved software interface between network layer and modem processing; (4) improved wake-up processing, i.e., detection of 2-of-3 or 3-of-4 tones, rather than 3-of-3; (5) improved acquisition signal, i.e., one long chirp, rather than three short chirps; (6) improved channel-estimation diagnostics; (7) logging of channel estimates; (8) RTS/CTS handshaking; (9) configurable enabling of RTS/CTS handshake; (10) power control; (11) watchdog; (12) ARQ; (13) packet time-stamping; and (14) a simple form of adaptive modulation restricted solely to parameter selection for Hadamard MFSK modulation.

The new ATM885 hardware and the *Seaweb 2000* protocols are major strides toward the ultimate goal of a self-configuring, wireless network of autonomous undersea devices.

CONCLUSION

Telesonar is an emerging technology for wireless digital communications in the undersea environment. Telesonar transmission channels include shallow-water environments with node-to-node separations hundreds of times greater than the water depth. Robust, environmentally adaptive acoustic links interconnect undersea assets, integrating them as a unified resource.

Seaweb offers a blueprint for telesonar network infrastructure. Warfare considerations stipulate that the network architecture will support rapid installation, wide-area coverage, long standoff range, invulnerability, and cross-mission interoperability. Seaweb is an information system compatible with low bandwidth, high latency, and variable quality of service. Seaweb connectivity emphasizes reliability, flexibility, affordability, energy efficiency, and transmission security. Network interfaces to manned command centers via gateway nodes such as the racom buoy are an essential aspect of the seaweb concept. C³N via seaweb supports common situational awareness and collective adaptation to evolving rules of engagement. Seaweb revolutionizes naval warfare by ultimately extending network-centric operations into the undersea battlespace.

The *Seaweb '98*, *'99*, and *2000* experiments incrementally advanced telesonar underwater acoustic signaling and ranging technology for undersea wireless networks. The constraints imposed by acoustic transmission through shallow-water channels have yielded channel-tolerant signaling methods, hybrid multi-user access strategies, novel network topologies, half-duplex handshake protocols, and iterative power-control techniques. *Seawebs '98* and *'99*, respectively, included 10 and 15 battery-powered, anchored telesonar nodes organized as non-centralized, bidirectional networks. These tests demonstrated the feasibility of battery-powered, wide-area undersea networks linked via radio gateway buoy to the terrestrial internet. Testing involved delivery of remotely sensed data from the sea and remote control from manned command centers ashore and afloat. *Seaweb 2000* included 17 nodes equipped with new telesonar modem hardware. It introduced a compact protocol anticipating adaptive network development.

In late summer, *Seaweb 2001* will be conducted in a large expanse of 30- to 300-m waters adjacent to San Diego, CA. The annual seaweb experiments will continue to extend area coverage, resource optimization, network capacity, functionality, and quality of service. Active research includes spread-spectrum signaling, directional transducers [31], *in situ* channel estimation, adaptive modulation, *ad hoc* network initialization, and node ranging and localization.

SSC San Diego is applying seaweb technology for ocean surveillance (DADS Demonstration Project), littoral ASW (Hydra Project), oceanographic research (FRONT Project), submarine communications (Sublink Project) and UUV command and control (SLOCUM and EMATT). Additional applications are proposed.

Deployable autonomous undersea systems will enhance the warfighting effectiveness of submarines, maritime patrol aircraft, amphibious forces, battle groups, and space satellites. Wide-area sensor grids, leave-behind multistatic sonar sources, mine-hunting robots, swimmer-delivery systems, and autonomous vehicles are just a few of the battery-powered, offboard devices that will augment high-value space and naval platforms. Distributed system architectures offer maximum flexibility for addressing a wide array of ocean environments and military missions.

ACKNOWLEDGMENTS

SSC San Diego established the Seaweb Initiative to advance science and technology capabilities for naval command, control, communications, computers, intelligence, surveillance, and reconnaissance (C⁴ISR).

Personnel contributing to the Seaweb Initiative are Keyko McDonald, Dave Rees, Dick Shockley, Mark Hatch, Tedd Wright, and Joan Kaina.

The primary sponsor of the Seaweb experiments is the ONR Sensors, Sources, and Arrays Program (ONR 321SS). ONR 321SS also funded the development of telesonar test beds, directional telesonar transducers, and the seaweb server. The SSC San Diego In-house Laboratory Independent Research (ILIR) Program sponsored the telesonar channel measurements and modeling reported here.

Benthos, Inc. personnel Ken Scussel, Dave Porta, John Baker, Jim Hardiman, Dale Green, Jack Crosby, Steve Niland, Bob Burns, Tom Tuite, and Steve Fantone contributed to this work. Benthos' telesonar products and their involvement in seaweb development resulted from the August 1999 acquisition of Datasonics, Inc.

Delphi Communication Systems (DCS) personnel Michael Wolf, Steve Merriam, Ethem Sozer, John Proakis, and Rami Mehio contributed to this work.

DCS and Datasonics involvement in *Seaweb '98* was sponsored by an SBIR topic N93-170 Phase-2 contract administered by Phil DePauk and Linda Whittington of SPAWAR. DCS and Benthos' involvement in *Seawebs '99* and *2000* was sponsored by an SBIR topic N97-106 Phase-2 contract administered by Al Benson, Tom Curtin, Don Davison, Dave Johnson, and Doug Harry of ONR. In addition, Datasonics and Benthos invested independent research and development (IR&D) resources in support of *Seawebs '98* and *'99*.

Dan Codiga, Philip Bogden, Dennis Arbige, and Adam Houk of the University of Connecticut's Department of Marine Sciences contributed to *Seawebs '99* and *2000* and implemented the cellular-modem gateway and the ADCP sensor node. John Newton (Polar Associates, Inc.) performed environmental analysis in support of *Seaweb '98*. Kent Raysin (SAIC) participated in *Seaweb '99*, and Mike Porter (SAIC) participated in the SignalEx portion of *Seaweb 2000*.

AUTHORS

Robert K. Creber

BS in Physics, Florida Institute of Technology, 1984
Current Research: Undersea acoustic modems and networks.

Christopher L. Fletcher

BS in Electrical Engineering, University of Illinois at Champaign/Urbana, 1997
Current Research: Telesonar modems; undersea acoustic networks; lightweight underwater acoustic arrays.

Paul A. Baxley

MS in Oceanography, University of California at San Diego, 1998
Current Research: Underwater acoustic communication modeling; matched-field source localization and tracking; seafloor geoacoustic property inversion.

Kenneth E. Rogers

MS in Electrical Engineering, Brigham Young University, 1969
Current Research: Undersea sensors and gateway communications for deployable surveillance systems.

Donald C. Davison

Ph.D. in High Energy Physics, University of California at Riverside, 1969
Current Research: Off-board and deployable surveillance systems.



Joseph A. Rice

M.S. in Electrical Engineering, University of California at San Diego, 1990
Current Research: Ocean sound propagation; sonar systems analysis; undersea wireless networks.

REFERENCES

1. Kilfoyle, D. B. and A. B. Baggeroer. 2000. "The State of the Art in Underwater Acoustic Telemetry," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 4–27.
2. Catipovic, J. A., M. Deffenbaugh, L. Freitag, and D. Frye. 1989. "An Acoustic Telemetry System for Deep Ocean Mooring Data Acquisition and Control," *Proceedings of IEEE Oceans '89 Conference*, September.
3. Merriam, S. and D. Porta. 1993. "DSP-Based Acoustic Telemetry Modems," *Sea Technology*, May.
4. Porta, D. 1996. "DSP-Based Acoustic Data Telemetry," *Sea Technology*, February.
5. Rice, J. A. and K. E. Rogers. 1996. "Directions in Littoral Undersea Wireless Telemetry," *Proceedings of The Technical Cooperation Program (TTCP) Symposium on Shallow-Water Undersea Warfare*, vol. 1, pp. 161–172.
6. Scussel, K. F., J. A. Rice, and S. Merriam. 1997. "New MFSK Acoustic Modem for Operation in Adverse Underwater Acoustic Channels," *Proceedings of IEEE Oceans '97 Conference*, Halifax, Nova Scotia, Canada, October, pp. 247–254.
7. Green, M. D. 2000. "New Innovations in Underwater Acoustic Communications," *Proceedings: Oceanology International*, March, Brighton, UK.
8. Jahn, E., M. Hatch, and J. Kaina. 1999. "Fusion of Multi-Sensor Information from an Autonomous Undersea Distributed Field of Sensors," *Proceedings of Fusion '99 Conference*, July, Sunnyvale, CA.
9. McGirr, S., K. Raysin, C. Ivancic, and C. Alspaugh. 1999. "Simulation of Underwater Sensor Networks," *Proceedings of IEEE Oceans '99 Conference*, September, Seattle, WA.
10. Curtin, T. B., J. G. Bellingham, J. Catipovic, and D. Webb. 1993. "Autonomous Oceanographic Sampling Networks," *Oceanography*, vol. 6, pp. 86–94.
11. Sozer, E. M., M. Stojanovic, and J. G. Proakis. 2000. "Underwater Acoustic Networks," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 72–83.
12. Rice, J. A. 2000. "Telesonar Signaling and Seaweb Underwater Wireless Networks," *Proceedings of NATO Symposium on New Information Processing Techniques for Military Systems*, 9 to 11 October, Istanbul, Turkey.
13. Rice, J. A. 1997. "Acoustic Signal Dispersion and Distortion by Shallow Undersea Transmission Channels," *Proceedings of NATO SACLANT Undersea Research Centre Conference on High-Frequency Acoustics in Shallow Water*, July, pp. 435–442.
14. Rice, J. A. and R. C. Shockley. 1998. "Battery-Energy Estimates for Telesonar Modems in a Notional Undersea Network," *Proceedings of the MTS Ocean Community Conference*, Marine Technical Society, vol. 2, pp. 1007–1015.
15. Karn, P. 1990. "MACA-A New Channel Access Method for Packet Radio," *Proceedings of the Amateur Radio Relay League/Canadian Radio Relay League (ARRL/CRRL) 9th Computer Network Conference*, September.
16. Rice, J. A. and M. D. Green. 1998. "Adaptive Modulation for Undersea Acoustic Modems," *Proceedings of the MTS Ocean Community Conference*, vol. 2, pp. 850–855.
17. Green, M. D. and J. A. Rice. 2000. "Channel-Tolerant FH-MFSK Acoustic Signaling for Undersea Communications and Networks," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 28–39.
18. Sozer, E. M., J. G. Proakis, M. Stojanovic, J. A. Rice, R. A. Benson, and M. Hatch. 1999. "Direct-Sequence Spread-Spectrum-Based Modem for Underwater Acoustic Communication and Channel Measurements," *Proceedings of IEEE Oceans '99 Conference*, September, Seattle, WA.

19. Stojanovic, M., J. G. Proakis, J. A. Rice, and M. D. Green. 1998. "Spread-Spectrum Methods for Underwater Acoustic Communications," *Proceedings of IEEE Oceans '98 Conference*, September, vol. 2, pp. 650–654.
20. Fruehauf, N. and J. A. Rice. 2000. "System Design Aspects of a Steerable Directional Acoustic Communications Transducer for Autonomous Undersea Systems," *Proceedings of Oceans 2000 Conference*, September, Providence, RI.
21. Rice, J. A., V. K. McDonald, M. D. Green, and D. Porta. 1999. "Adaptive Modulation for Undersea Acoustic Telemetry," *Sea Technology*, vol. 40, no. 5, pp. 29–36.
22. Raysin, K., J. A. Rice, E. Dorman, and S. Matheny. 1999. "Telesonar Network Modeling and Simulation," *Proceedings of IEEE Oceans '99 Conference*, September.
23. Proakis, J. G., M. Stojanovic, and J. A. Rice. 1998. "Design of a Communication Network for Shallow-Water Acoustic Modems," *Proceedings of MTS Ocean Community Conference*, November, vol. 2, pp. 1150–1159.
24. McDonald, V. K., J. A. Rice, M. B. Porter, and P. A. Baxley. 1999. "Performance Measurements of a Diverse Collection of Undersea Acoustic Communication Signals," *Proceedings of IEEE Oceans '99 Conference*, September, Seattle, WA.
25. Codiga, D. L., J. A. Rice, and P. S. Bogden. 2000. "Real-Time Delivery of Subsurface Coastal Circulation Measurements from Distributed Instruments Using Networked Acoustic Modems," *Proceedings of IEEE Oceans 2000 Conference*, September, Providence, RI.
26. Baxley, P. A., H. P. Buckner, and J. A. Rice. 1998. "Shallow-Water Acoustic Communications Channel Modeling Using Three-Dimensional Gaussian Beams," *Proceedings of MTS Ocean Community Conference*, vol. 2, pp. 1022–1026.
27. Green, M. D., J. A. Rice, and S. Merriam. 1998. "Underwater Acoustic Modem Configured for Use in a Local Area Network," *Proceedings of IEEE Oceans '98 Conference*, September, vol. 2, pp. 634–638.
28. Green, M. D., J. A. Rice, and S. Merriam. 1998. "Implementing an Undersea Wireless Network Using COTS Acoustic Modems," *Proceedings of MTS Ocean Community Conference*, vol. 2, pp. 1027–1031.
29. McDonald, V. K. and J. A. Rice. 1999. "Telesonar Testbed Advances in Undersea Wireless Communications," *Sea Technology*, vol. 40, no. 2, pp. 17–23.
30. Porter, M. B., V. K. McDonald, J. A. Rice, and P. A. Baxley. 2000. "Relating the Channel to Acoustic Modem Performance," *Proceedings of the European Conference on Underwater Acoustics*, July, Lyons, France.
31. Butler, A. L., J. L. Butler, W. L. Dalton, and J. A. Rice. 2000. "Multimode Directional Telesonar Transducer," *Proceedings of IEEE Oceans 2000 Conference*, September, Providence, RI.



Shallow-Water Acoustic Communications Channel Modeling Using Three-Dimensional Gaussian Beams

Paul A. Baxley, Homer Buckner, Vincent K. McDonald,
and Joseph A. Rice
SSC San Diego

Michael B. Porter
SAIC/Scripps Institution of Oceanography

INTRODUCTION

Recent innovations in shallow-water undersea surveillance and exploration have necessitated the use of the underwater acoustic medium as the primary means of information exchange. Wireless communication between underwater stations separated in range by as much as 5 km with water depths as low as 10 m may be required. This task is complicated by the inherent spatiotemporal variability of this medium, and the complex nature of multipath arrival of energy for shallow-water environments [1]. Figure 1 illustrates some of the major processes that may affect underwater communication signals.

Multipath spread is caused by refraction governed by the sound-speed profile, reflections from boundaries, and scattering from inhomogeneities. Doppler spread arises from source/receiver motion or the motion of the reflectors and scatterers. These phenomena can significantly disperse and distort the signal as it propagates through the channel. A numerical propagation model that simulates these effects is desired for the systematic study of these phenomena. Such a model would also be a useful tool for environment-dependence assessment, performance prediction, and mission planning of communication systems.

This paper describes an approach for a physics-based model designed to simulate multipath spread and Doppler spread of high-frequency underwater acoustic communication signals.

Multipath spread is handled via propagation through a refractive medium, as dictated by the sound-speed profile, and by the modeling of reflection and scattering from arbitrarily rough boundaries. Doppler spread is incorporated via the inclusion of source/receiver motion and sea-surface motion. While other phenomena (water mass fluctuations, scattering from water volume inhomogeneities or bubbles) can be responsible for signal distortions, it is believed that those included are the primary sources of spreading for

ABSTRACT

Recent progress in the development of a physics-based numerical propagation model for the virtual transmission of acoustic communication signals in shallow water is presented. The ultimate objective is to provide for the prediction of the output of the quadrature detector (QD, an analog of the discrete Fourier transform) in a time-variant, doubly dispersive, shallow-water channel. Current model development concentrates on the modeling of the QD response in the presence of rough boundaries, reserving inclusion of effects caused by a time-varying sea surface or source/receiver motion to future implementations. Three-dimensional Gaussian beam tracing is used so that out-of-plane reflections from rough surfaces or sloping bathymetry can be adequately modeled. Model predictions of the impulse response for a real shallow-water environment are observed to agree well with measured impulse responses.

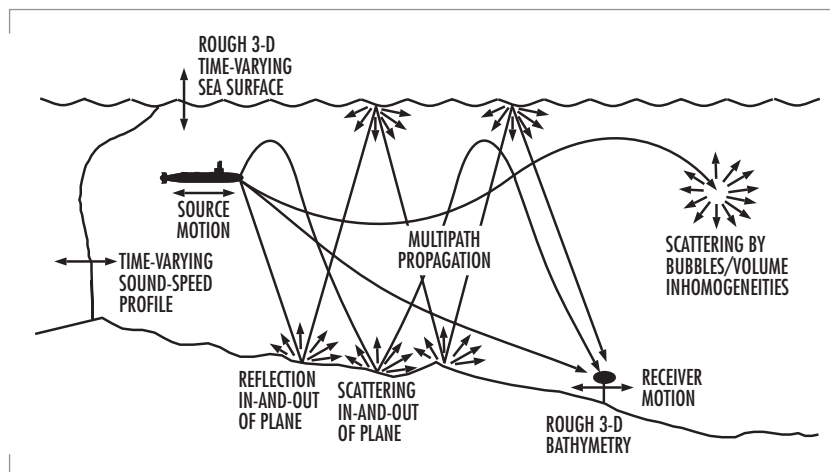


FIGURE 1. Some of the major processes affecting underwater acoustic communications signals.

many realistic problems. The present emphasis is on the modeling of propagation in a bounded refractive medium and the scattering from rough boundaries. The effects of a time-varying sea surface and source/receiver motion are reserved for future implementations.

The basic approach is to model the received output of the quadrature detector (QD) for a transmitted finite-duration constant-wavelength (CW) pulse. The QD is an analog version of the discrete Fourier transform, and provides a convenient means of obtaining the complex Fourier coefficients as a function of time for a finite-duration CW pulse. Because finite-duration CW pulses are common signals in communication schemes, the modeling of such signals is appropriate. However, a broadband QD response (for multiple CW pulses of different frequencies) can also be used to obtain a band-limited impulse response via Fourier synthesis, which is useful for the study of any arbitrary pulse signature.

The pulse is propagated by means of three-dimensional (3-D) Gaussian beams. The consideration of propagation in three dimensions is important because energy can be reflected or scattered in and out of the vertical plane containing both the source and receiver. The high frequencies of communication signals dictate the use of ray-based models over the less-efficient wave models or parabolic-equation approximations. Ray-based models also ensure proper handling of range-dependence and proper reflections from sloping boundaries. The only ray-based method practical for the 3-D problem is Gaussian beams, because the necessity of eigen-ray determination is eliminated. Ray theory without the use of Gaussian beams requires the determination of eigenrays (rays following paths connecting the source and receiver exactly), which is a formidable task in three dimensions. A dense fan of Gaussian microbeams allows direct modeling of scattering from arbitrarily rough surfaces.

EXAMPLE OF THE EFFECT OF PROPAGATION CONDITIONS ON COMMUNICATIONS

A compelling example of how ocean channel physics can affect underwater communications was

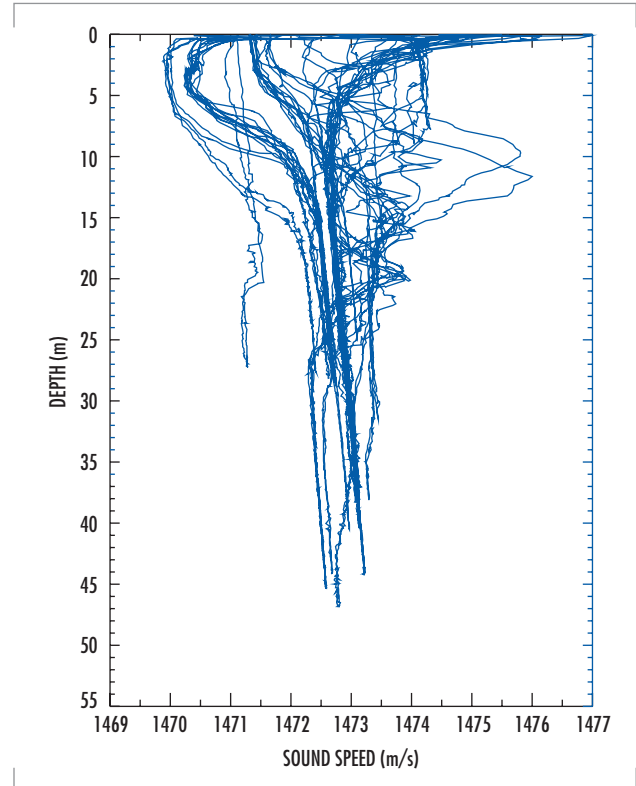


FIGURE 2. Composite of sound-speed profiles measured during the FRONT engineering test.

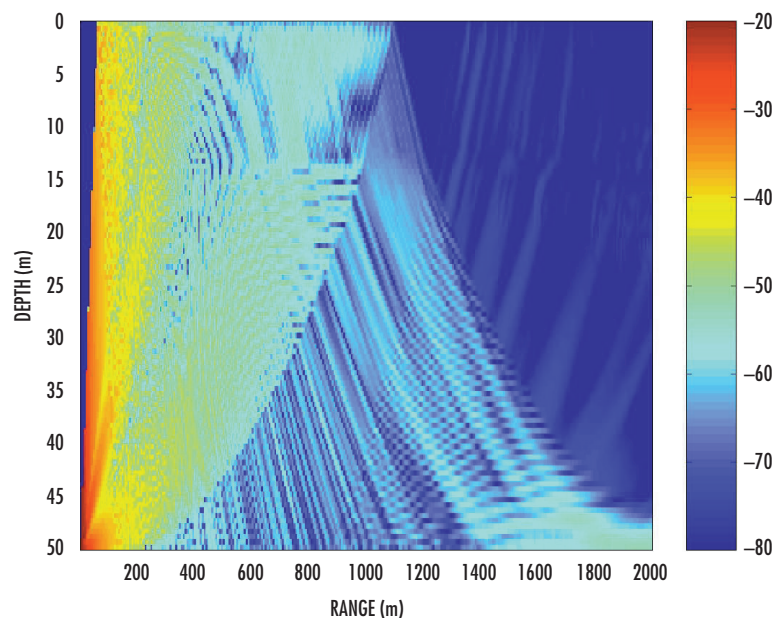


FIGURE 3. Predicted transmission loss during the FRONT engineering test.

provided in engineering tests for the Front-Resolving Observatory with Networked Telemetry (FRONT) oceanographic network. The oceanographic conditions in the area are both interesting and complicated as fresh-river runoff interacts with the tides to generate a persistent front. Figure 2 shows a composite of sound-speed profiles measured at various locations over the duration of the experiment, demonstrating the great variability in the region. Within the upper 10 m of the water column, the channel varies between an upward-refracting (sound speed increases with depth) and downward-refracting (sound speed decreases with depth) channel. Figure 3 shows a transmission loss plot for a typical upward-refracting profile and a communications node (serving as the projector) located on the ocean bottom. This suggests that the influence of sea-surface roughness and time-variability will be greater when the channel is upward refracting.

During the course of the network deployment, there were periods with strong winds followed by relatively calm conditions as the wind speed plot shows in Figure 4A. As the wind speed increases, wave action drives up the ambient noise. At the same time, the roughness of the surface makes it a poor acoustic reflector, so the signal level drops. The combination of the two factors drives the signal-to-noise ratio (SNR) at the bottom-mounted receiver (Figure 4B). This, in turn, drives the overall modem performance as measured by the bit-error rate (Figure 4C). In summary, high winds caused network outages.

This is the simplest of mechanisms driving modem performance. Even with strong SNR, a modem that relies on a tap-delay line for adaptive equalization may fail if the multipath spread becomes too long. Similarly, a modem may fail to track Doppler changes, which is yet another dimension to the parameter space affecting modem performance.

3-D GAUSSIAN BEAM PROPAGATION MODEL

The 3-D Gaussian beam model is a modified version of that presented by Bucker [2]. For a specified sound-speed profile and seafloor, beams are traced from a source in three dimensions following the laws of ray acoustics and boundary interactions. Ray theory requires the determination of eigenrays, which can be computationally intensive, particularly in three dimensions. Determination of eigenrays is unnecessary in the Gaussian beam formulation. The sound field at a receiver is obtained by combining the coherent contributions of each beam as determined by the closest point of approach (CPA) of the beam path to the receiver. Consider an arbitrary beam path p that travels from a source S and passes close to a receiver X , as shown schematically in Figure 5A. The point x represents the CPA of the beam to the receiver and ρ is the CPA distance. The actual path length (arc length) from

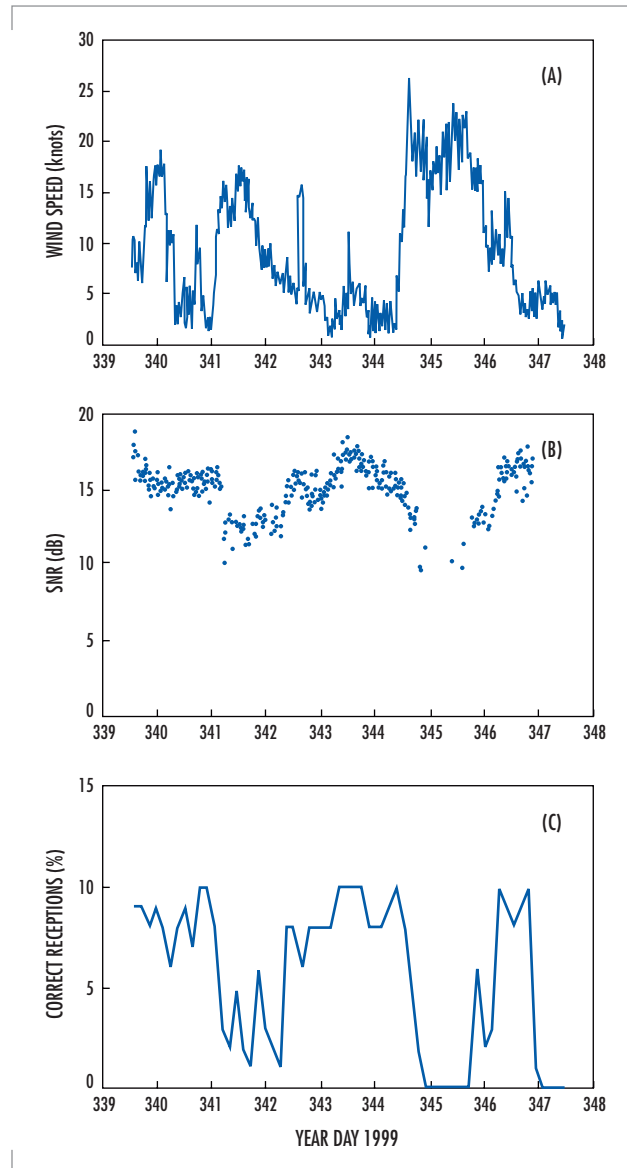


FIGURE 4. (A) During the FRONT network-engineering test, the wind speed varied considerably. (B) The wind speed affects the ambient noise and the surface reflectivity, which both drive the SNR at the receiver. (C) Variations in SNR, in turn, drive the performance of the modem.

S to x is designated S_x , and is shown linearized in Figure 5B. The pressure at the receiver X associated with this beam path is then given by

$$p = C_n B [\exp(-a\theta^2 + i\omega t)] / S_x, \quad (1)$$

where C_n is a normalization constant, a is an empirical constant, $\theta = \tan^{-1}(\rho / S_x)$, ω is the angular frequency, and t is the travel time to point x. A spherical wave-front correction equal to $(L - S_x)/c_x$, where $L^2 = S_x^2 + \rho^2$ and c_x is the sound speed at x, is included in the travel time t . B accounts for energy loss and phase shifts resulting from surface and bottom reflections. See [2] for a fuller explanation of the constants a and C_n .

An important feature of the 3-D Gaussian beam model is that the bottom can be specified arbitrarily. Bottom depth data z are specified digitally by the user as a function of the horizontal coordinate directions x and y . Third-order smoothing polynomials are fitted to the bottom data in both directions so that the depth z and the unit normal \tilde{n} can be determined for any arbitrary value of x and y (see appendix B of [2] for details). Bottom interactions are modeled via the specification of the reflection coefficient, or via the calculation of the reflection coefficient from specified geoacoustic parameters (sediment compressional and shear sound speed and attenuation and density). The bottom displacement technique of Zhang and Tindle [3] may also be used. Currently, only a semi-infinite representation for the bottom is implemented; this is sufficient for the high frequencies of interest in communication systems.

Arbitrary specification of the bottom depth implies that scattering problems may be handled directly and deterministically without the use of statistical techniques or approaches only applicable to particular classes of problems because of their underlying assumptions. Because the roughness can be arbitrarily specified, scattering effects are modeled by tracing a dense fan of very fine microbeams, which follow the physics of the interface interactions directly. By this means, problems at shallow grazing angles, such as self-shadowing effects, can be treated. In addition, using an arbitrary specification of bottom depth with 3-D Gaussian beams allows examination of environments possessing significant range-dependence.

QUADRATURE DETECTOR RESPONSE

Assume that a continuous sinusoidal signal $A \cos(\omega t + \phi)$ arrives at a receiver, where ω is the frequency, t is the time, and ϕ is a phase shift associated with boundary interactions. If the signal is processed by a quadrature detector, as diagramed in Figure 6, the signal is split with one part being multiplied by $2 \cos(\omega_0 t)$ and the other part being multiplied by $-2 \sin(\omega_0 t)$. ω_0 is the reference frequency, which should be approximately equal to ω . Both parts are then passed through a low-pass filter to obtain the quadrature components $A \cos[(\omega - \omega_0)t + \phi]$ and $A \sin[(\omega - \omega_0)t + \phi]$, respectively. The complex output R_{QD} of the QD is then simply

$$R_{QD} = A \exp[i(\omega - \omega_0)t + \phi], \quad (2)$$

This is basically an analog version of the discrete Fourier transform. If $\omega_0 \neq \omega$, R_{QD} will experience a rotation of $\omega - \omega_0$ radians per second.

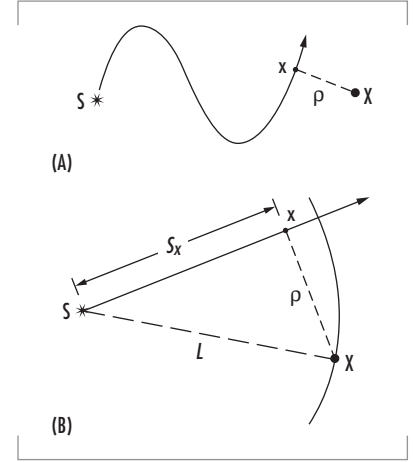


FIGURE 5. Geometry used to determine pressure contribution of a Gaussian beam at CPA to sensor X.

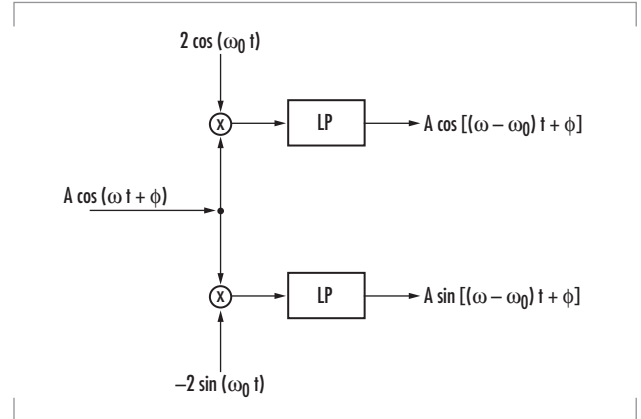


FIGURE 6. Quadrature detector algorithm for a continuous sinusoidal signal.

Now assume that the incoming signal is a finite sinusoidal pulse of duration τ seconds and frequency ω . Assume also that the travel time from the source to the receiver along path p is t_p and that the time constant t_c (effective integration time) of the low-pass filters in the QD is τ seconds. For this case, R_{QD} is modulated by a triangle function $T(t)$ that is zero for $t < t_p$, increases linearly from zero to a value of unity at $t = t_p + \tau$, and then decreases linearly back to zero at $t = t_p + 2\tau$. Therefore, the quadrature response for a beam travelling along path p is

$$R_{QD} = AT(t)\exp[i(\omega - \omega_0)t + \phi], \quad (3)$$

If the time constant t_c is larger than the pulse duration τ , R_{QD} remains at the value of the apex until $t > t_p + t_c$. In either case, R_{QD} still experiences the rotation $\omega - \omega_0$ if $\omega_0 \neq \omega$. The only way that ω_0 cannot equal ω in the above scenario is for a Doppler shift to have occurred somewhere along the path p . Therefore, source/receiver motion or sea-surface motion results in a rotation of R_{QD} for a path influenced by that motion.

The total quadrature response of a received signal is therefore easily obtained by combining the quadrature responses of all paths contributing to the pressure at the receiver. This is facilitated via the use of the 3-D Gaussian beam model to propagate the energy. Closely spaced microbeams are launched from the source and traced through the refractive, bounded medium. Travel times, phase shifts associated with boundary reflections, and Doppler shifts associated with a moving source/receiver or with a moving sea surface are accumulated for each microbeam as it propagates. This information is then used with Eq. (3) to determine the QD response for each microbeam. A superposition of the QD response for all microbeams then provides the total QD response. If the pulse length is small, the QD response represents an estimate of the channel impulse response. The multipath structure resulting from refraction and the complex interactions of the many microbeams with the rough surface will combine to yield the effect of multipath spread on the QD impulse response. The Doppler shifts accumulated for each microbeam will combine to yield the effect of Doppler spread on the QD impulse response.

MODEL DEMONSTRATION

The environment selected for demonstrating the usefulness of the channel model was that of the SignalEx-99 experiment conducted in April 1999 in a shallow-water (~ 200 m) region, 6 km southwest of San Diego. Sponsored by the Office of Naval Research, SignalEx-99 was the first in a series of experiments intended to relate channel propagation characteristics to the performance of underwater acoustic communication systems. A detailed description of the experiment is provided by McDonald et al. [4].

The data considered here were linear frequency-modulated (LFM) chirps emitted/received from a source/receiver deployed at a depth of 30 m and source/receiver mounted 6.7 m above the seafloor. The source/receiver systems were telesonar test beds [5 and 6], autonomous units consisting of a single-board computer with a projector and a four-phone vertical line array. The 30-m test bed was deployed from a freely drifting ship, resulting in measurements as a function of time along a fairly constant track. The water depth at the receiver was 210 m, while the water depth decreased in a near linear fashion along the track to a depth of approximately 170 m at a range of 3.8 km from the receiver. Transmissions were made in both directions between the two test beds.

Figure 7 shows the bathymetry and track of the drifting source. Northerly winds caused the ship to drift from a range of about 0 to 4 km (Drift 1). As the range was becoming large and the ship began drifting off the isobath, the ship was repositioned back at a range of about 2 km and allowed to drift again (Drift 2). This conveniently provided a look at the consistency of the Drift 1 results. Once again, the ship drifted to a range of about 6.5 km and was repositioned and moored at a range of 4.75 km providing a look at the stability of the signaling schemes with fixed source-receiver geometry.

Figure 8 shows a typical sound-speed profile measured during the SignalEx-99 experiment. The profile is strongly downward refracting with approximately a 20-m mixed layer at the surface and a slight duct forming near the bottom. It has been determined previously [7] that the bottom in this region may be treated as a fluid with a compressional sound speed of 1572.37 m/s, a compressional attenuation of 0.20 dB/kmHz, and a density of 1.76 g/cm³.

The LFM chirps were transmitted sweeping the 8- to 16-kHz band over a 1-second period. Sixteen chirps were transmitted in 10-minute time frames over a 5-hour period. The direction of the transmission was switched for consecutive 10-minute periods. Theoretically, the impulse response is a combination of these chirps delayed in time according to their path length and attenuated according to volume absorption and reflection loss at the boundaries. The impulse response can be estimated experimentally by correlating the received pulses with a replica of the original transmitted pulse. This produces a sequence of impulses corresponding to each echo in the received waveform, thereby providing a visualization of the impulse response. Figure 9 shows the result of performing this correlation as a function of time. The variation in the multipath structure throughout the experiment is clearly observed. Because absolute times were not available, the first significant peak in each reception was detected and used to provide a leading-edge alignment. Note also that this plot is a composite of the transmissions that alternated between the ship and bottom-mounted test bed.

Figure 10 compares the measured impulse response at a time of day of 12.5 hours (ping number 34) to a simulated response obtained via the 3-D Gaussian beam, quadrature detector model for the same source-receiver configuration. The source range at this time was 2.2 km. The simulation was performed assuming a flat bottom at a water depth of 210 m, and ignoring the effects of rough-surface scattering and time-variability. The measured impulse response has been normalized relative to the maximum, and the arbitrary time scale has been shifted to facilitate comparison with the modeled result. Note that the predicted arrivals agree well with the measured arrivals, indicating that refraction and reflection

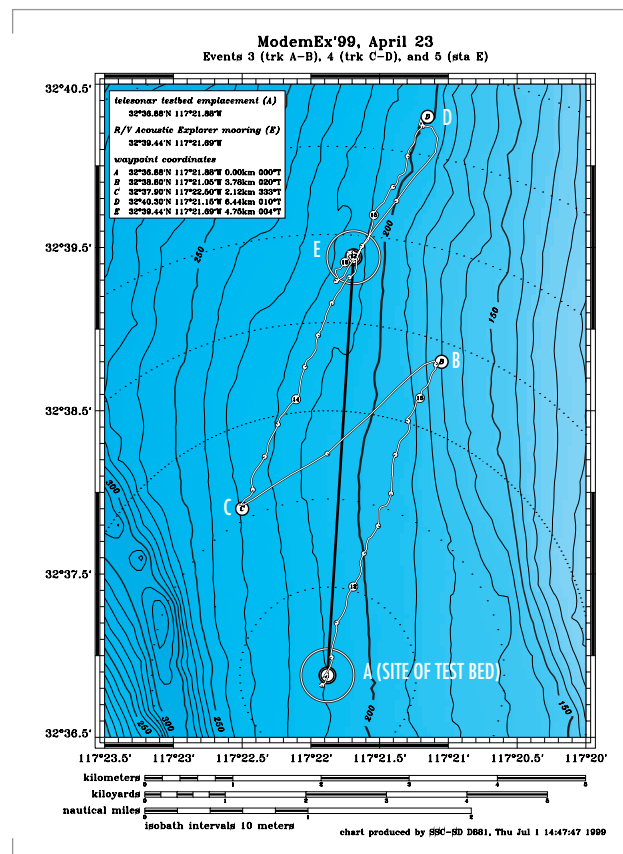


FIGURE 7. Bathymetry and source track for SignalEx-99 experiment. Drift 1 is from A to B (source range = 0 to 3.8 km). Drift 2 is from C to D (source range = 2.1 to 6.5 km). The moored station is at E (source range = 4.75 km).

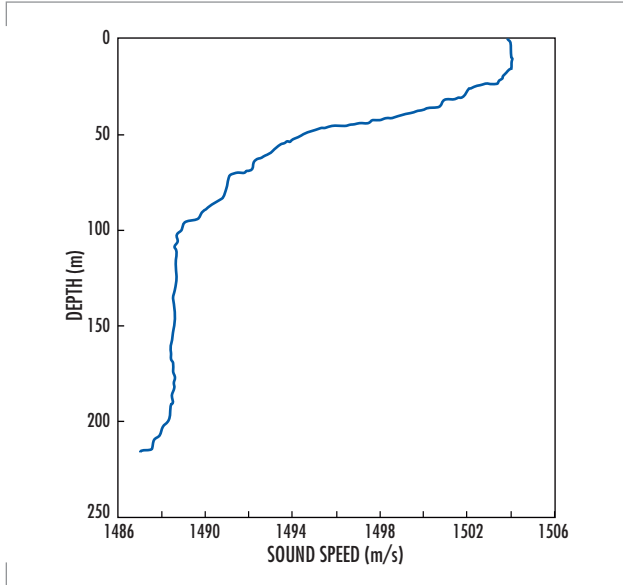


FIGURE 8. Typical sound-speed profile measured during SignalEx-99 experiment.

from boundaries are well modeled. Time discrepancies between arrival paths may be caused by the neglect of the varying bathymetry or errors in the assumed sound-speed profile. The higher resolution of the model results indicates that the first arrival is actually a combination of several arrivals: namely, the direct path, the one-bottom-reflected path, the one-surface-reflected path, and the one-surface-reflected/one-bottom-reflected path. Likewise, the later arrivals are actually a combination of several higher order paths. Note also that the data exhibits a gradual rolloff after the arrival of the pulses, suggesting a reverberant environment. The likely cause of this behavior is the scattering of energy in three dimensions caused by the interaction of rays with the boundaries. Future work will attempt to model these interactions.

Future developments of the model will focus on determining how scattering from rough surfaces, source/receiver motion, and sea-surface motion will influence these responses.

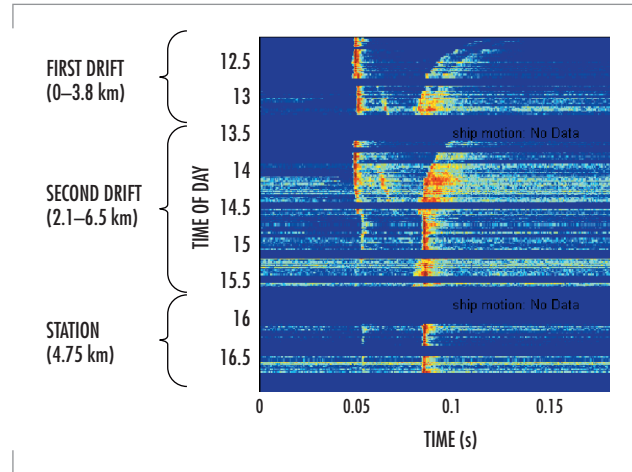


FIGURE 9. Replica correlogram from chirps during SignalEx-99 experiment.

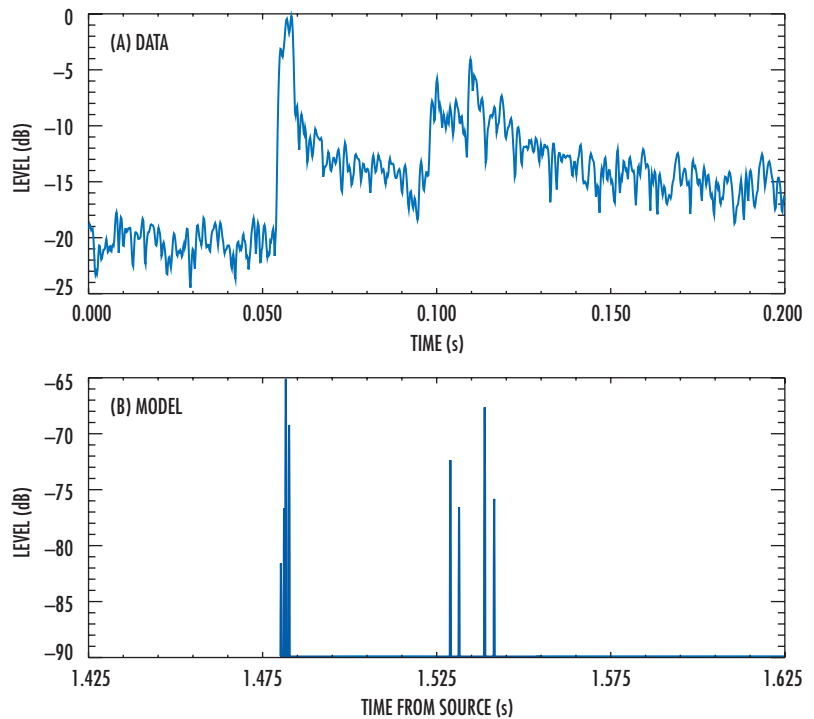


FIGURE 10. Comparison between (A) measured and (B) modeled impulse responses. Time of day = 12.5 hours. Source range = 2.2 km.

MODELING SURFACE AND BOTTOM SCATTER

Scattering from rough boundaries produces losses in signal energy. These losses are two-fold. First, scatter converts energy to higher angles eventually allowing it to penetrate the bottom where it is absorbed. Second, it destroys the coherence of the wave producing what might be termed an apparent loss. For instance, a moving surface will stretch and compress a sinewave reflected from it. If the reflected energy is detected by a matched-filter expecting a perfect sinewave, it will see a reduced power level. This discussion applies, for instance, to a single tone in an M-ary Frequency-Shift Keying (MFSK) signaling scheme, where the tone is detected by a filter bank. If we have a rough bottom with a static geometry, this loss of coherence does not occur. However, if the source or receiver moves, we have a dynamic situation similar to the surface loss just described.

In round numbers, a typical communications carrier gives a wavelength around 10 cm. A classical measure of the role of roughness—the Rayleigh roughness parameter—is the ratio of the roughness to the wavelength (or more precisely, the vertical component of the wavelength). As this number becomes close to unity, losses per bounce become large, perhaps 10 dB, and many of the standard scatter models that assume small roughness fail. The point of this discussion is that 10-cm roughness is easily attained on both surface and bottom boundaries in real environments, implying large boundary losses. Furthermore, the roughness is typically not known to within 10 cm, implying large uncertainty in those same losses and in the resulting transmission loss.

Finally, the actual scatter mechanisms are complicated. In some cases, the air–water interface is the scatterer. In other cases, the bubbles below are likely to be dominant. Similarly, at the ocean bottom, scatter can occur at the interface or by inhomogeneities just below the interface (though not too far below because volume attenuation limits the sediment penetration significantly).

As a first step toward modeling scattering effects, we assume that the boundary roughness dominates the problem, and concentrate first on the bottom roughness. A common approach [8] to characterize this roughness is to use the spatial power spectral density, i.e., the power spectrum of the bottom roughness. Various forms may be used; however, one popular choice is $\Phi(k) \propto k^{-b}$, where b is a measured parameter for the particular site. Suggested values for b are given in [8] along with the RMS roughness that defines the overall amplitude of the spectrum.

Given the spatial power spectral density, we can construct individual realizations of the bottom by using a standard technique. In particular, we convert the power spectrum to an amplitude spectrum by taking its square root. We discretely sample the amplitude and then introduce a random phase. Finally, we do a fast Fourier transform (FFT) to produce and add in the mean depth to obtain a single realization of the bottom. In equations:

$$D(r) = \int A(k) e^{i\theta} e^{ikr} dk \quad (4)$$

where $A(k) = \sqrt{\Phi(k)}$ and $\Phi(k) = 5.5 \times 10^{-5} k^{-2.25}$ is the spatial power density spectrum for a particular area.

As a specific example, Figure 11 shows a single realization of the bottom depth using the above described power spectrum. Figure 12 compares the

predicted transmission loss for the SignalEx-99 environment using this rough bottom (Figure 12A) with that predicted using a smooth bottom (Figure 12B). The transmission loss calculation was done using the BELLHOP ray/beam model [9], which is a two-dimensional (2-D) version of the 3-D Gaussian beam model. The prediction is for the case of a test bed deployed on the seafloor. Note the fill-in of the shadow zone near the surface at a range of 2 km. There are also changes in the Lloyd-mirror pattern emanating from the source.

The root-mean-square (RMS) height used here is 0.23 m, which is fairly low. There will also be surface scatter that may also be expected to have a larger RMS roughness.

SUMMARY AND FUTURE APPLICATIONS

The channel model outlined in this paper is being developed to aid in the analysis of future underwater acoustic communication systems. The modeling of the QD response via the use of 3-D Gaussian beams enables the inclusion of physical phenomena known to influence such systems in a computationally efficient manner. The model will provide a useful tool for examining the effect of multipath and Doppler spread on the performance of these systems. Because it is designed for use with finite-duration CW pulses, the model can be used directly for the analysis of MFSK systems. Otherwise, the model can also be run for multiple frequencies to obtain a band-limited impulse response via Fourier synthesis.

The following future work is planned. An analytical model will be implemented for a sinusoidally

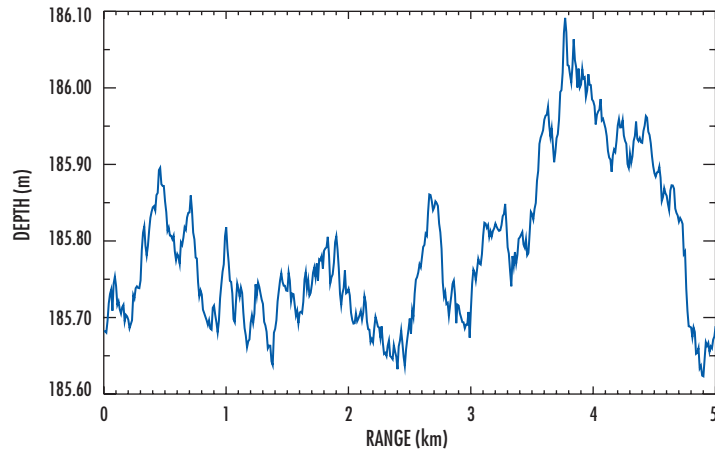


FIGURE 11. Single realization of the bottom depth using a power-law spectrum.

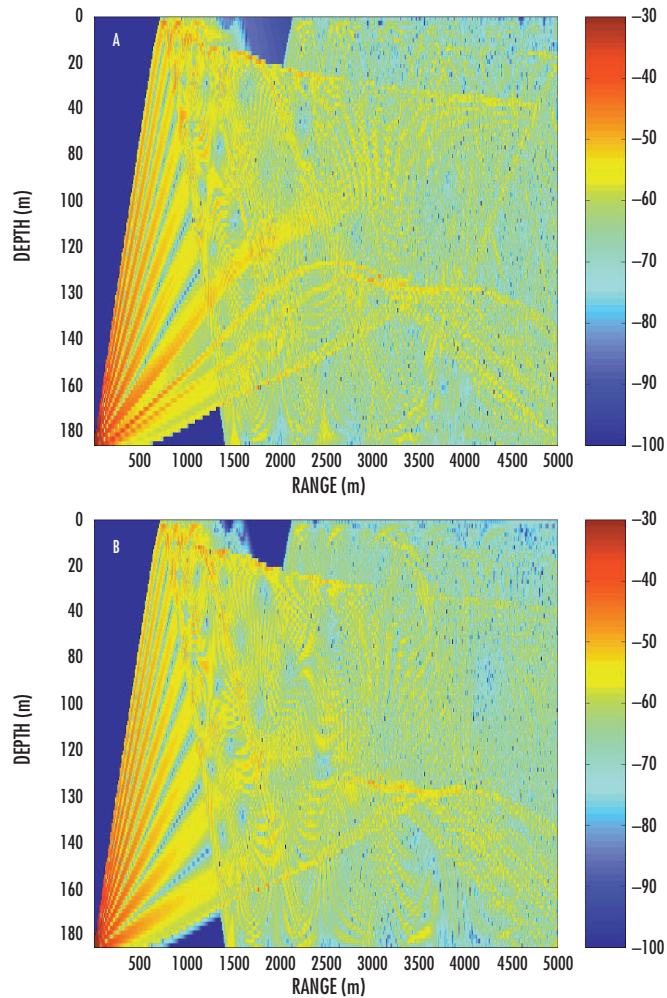


FIGURE 12. Comparison of predicted transmission loss for the SignalEx-99 environment (A) using the rough bottom in Figure 11, and (B) using a smooth bottom. Source is near the seafloor.

corrugated interface and used as a benchmark solution to validate the above ray/beam results for scatter from a rough surface. The 2-D scattering approach will be expanded to provide 3-D scattering for inclusion in the 3-D Gaussian beam quadrature detector model. The importance of 3-D scattering effects on the impulse response will then be studied. These tasks are geared to evaluating the mean energy level in a situation where the processing time is short enough that a "frozen ocean" model is appropriate. Once this is accomplished, a time-varying sea surface will be implemented.

Work is underway to use the model to infer channel characteristics such as coherence time, coherence bandwidth, multipath spread, and Doppler spread. It will also be used as part of a statistically governed Markov process to produce a time-dependent simulation.

ACKNOWLEDGMENTS

The Office of Naval Research (ONR) via the SSC San Diego In-house Laboratory Independent Research (ILIR) Program sponsored this work. The SignalEx-99 experiment was sponsored by ONR 322OM via the Telesonar Signaling Measurements Project. The FRONT program is sponsored by the National Oceanographic Partnership Program (NOPP). The authors also thank Robert Creber and Christopher Fletcher for their help in obtaining experimental data.

AUTHORS

Homer Bucker

Ph.D. in Physics, University of Oklahoma, 1962

Current Research: Matched-field tracking; underwater acoustic propagation modeling using three-dimensional Gaussian beams.

Vincent K. McDonald

BS in Mathematics, San Diego State University, 1988

Current Research: Underwater acoustic communications research; underwater surveillance system design.

Joseph A. Rice

MS in Electrical Engineering, University of California at San Diego, 1990

Current Research: Ocean sound propagation; sonar systems analysis; undersea wireless networks.

Michael B. Porter

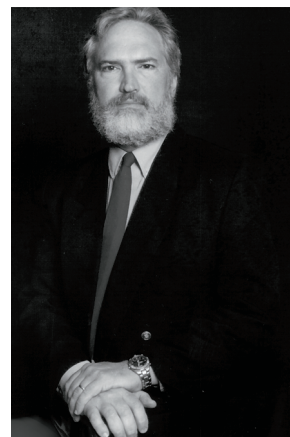
Ph.D. in Engineering Sciences and Applied Mathematics, Northwestern University, 1984

Current Research: Array signal processing; acoustic communications; wave propagation.

REFERENCES

1. Rice, J. A. 1997. "Acoustic Signal Dispersion and Distortion by Shallow Undersea Transmission Channels," *Proceedings of the NATO SACLANT Undersea Research Centre Conference on High-Frequency Acoustics in Shallow Water*, pp. 425–442.
2. Bucker, H. P. 1994. "A Simple 3-D Gaussian Beam Sound Propagation Model for Shallow Water," *Journal of Acoustical Society of America*, vol. 95, no. 5, pp. 2437–2440.

3. Zhang, Z. Y. and C. T. Tindle. 1993. "Complex Effective Depth of the Ocean Bottom," *Journal of Acoustical Society of America*, vol. 93, no. 1, pp. 205–213.
4. McDonald, V. K., J. A. Rice, M. B. Porter and P. A. Baxley. 1999. "Performance Measurements of a Diverse Collection of Undersea Acoustic Communication Signals," *Proceedings of IEEE Oceans '99 Conference*, 13 to 16 September, Seattle, WA.
5. McDonald, V. K., and J. A. Rice. 1999. "Telesonar Testbed—Advances in Undersea Wireless Communications," *Sea Technology*, vol. 40, no. 2, pp. 17–23.
6. McDonald, V. K., J. A. Rice, and C. L. Fletcher. 1998. "An Underwater Communication Testbed for Telesonar RDT&E," *Proceedings of MTS Ocean Community Conference '98*, 16 to 19 November, Baltimore, MD.
7. Baxley, P. A., N. O. Booth, and W. S. Hodgkiss. 2000. "Matched-Field Replica Model Optimization and Bottom Property Inversion in Shallow Water," *Journal of Acoustical Society of America*, vol. 107, no. 3, pp. 1301–1323.
8. Medwin, H. and C. S. Clay. 1998. *Fundamentals of Acoustical Oceanography*, Academic Press, San Diego, CA.
9. Porter, M. B. and Y-C Liu. 1994. "Finite-Element Ray Tracing," *Theoretical and Computational Acoustics*, (D. Lee and M. H. Schultz, eds.) World Scientific Publishing Company, River Edge, NJ, vol. 2, pp. 947–956.



Paul Baxley

MS in Oceanography, Scripps Institution of Oceanography, University of California at San Diego, 1998

Current Research: Underwater acoustic communication modeling; matched-field source localization and tracking; seafloor geoacoustic property inversion.

Advanced Refractive Effects Prediction System (AREPS)

Wayne L. Patterson
SSC San Diego

INTRODUCTION

In 1987, SSC San Diego provided the U.S. Navy's operational fleet with its first capability to assess the effects of the atmosphere on the performance of electromagnetic (EM) systems such as radars and radios. This assessment system was named the Integrated Refractive Effects Prediction System (IREPS). IREPS was hosted on the Hewlett-Packard 9845 desktop calculator. The EM propagation models of IREPS were semi-empirical and assumed that the atmosphere is homogeneous in the horizontal. IREPS also assumed the earth's surface was water. As desktop computing developed and EM propagation modeling advanced, the various assumptions of IREPS were overcome. In response to a request from Commander, Sixth Fleet during the Bosnian campaign, a new assessment system, the Advanced Refractive Effects Prediction System (AREPS) was fielded for fleet operations.

AREPS computes and displays radar probability of detection, propagation loss and signal-to-noise ratios, electronic-support-measures (ESM) vulnerability, UHF/VHF communications, and surface-borne surface-search radar capability vs. range, height, and bearing from the transmitter.

The power of AREPS derives from its Windows 95/NT interface, making full use of pop-up menus, object linking and embedding (OLE) features such as file drag and drop and graphics export, and extensive online help with color graphic examples.

At the core of AREPS is our Advanced Propagation Model (APM), a hybrid ray-optic and parabolic equation (PE) model that uses the complementary strengths of both methods to construct a fast yet very accurate composite model. Depending on the requirements of the tactical decision aid, APM will run in several different modes. For the full hybrid mode, APM is much faster than PE models alone, with overall accuracy at least as good as the pure PE models. With its airborne submodel, APM can solve problems for very high elevation angles where PE methods would not normally be used.

APM allows for range-dependent refractivity over various sea and/or terrain paths. Not only does the terrain path include variable terrain heights, it may also include range-varying dielectric ground constants for finite conductivity and vertical polarization calculations. APM considers absorption of electromagnetic energy by oxygen and water vapor. APM

ABSTRACT

In 1987, SSC San Diego fielded the Integrated Refractive Effects Prediction System (IREPS), the world's first electromagnetic prediction system for shipboard use. Advances in research and technology have led to the replacement of IREPS with the Advanced Refractive Effects Prediction System (AREPS). AREPS computes and displays radar probability of detection, propagation loss and signal-to-noise ratios, electronic-support-measures vulnerability, UHF/VHF communications, and surface-borne surface-search radar capability vs. range, height, and bearing from the transmitter.

accounts for all normal propagation mechanisms, including troposcatter and the anomalous propagation mechanisms of subrefraction, super-refraction, and ducting.

AREPS DISPLAYS

The primary AREPS displays are height vs. range and bearing coverage and path loss vs. height/range and bearing. Figure 1 shows such a coverage display for shipborne air-search radar with its probability of detecting a "small-sized" jet. For this case, the atmosphere is range-dependent, with a surface-based duct existing at the transmitter location, rising to become an elevated duct over the terrain features. To the lower right of the coverage display is a small map, in a simulated plan-position-indicator (ppi) picture format, showing the transmitter location, the display's current bearing, and the terrain heights.

At the top of the display window is a series of buttons that allow you to animate the display in bearing, both forward and backward, to pause the animation, and to obtain a printed copy of the display. Because AREPS is a Windows 95/NT program, the full capabilities of the operating system are available. For example, should you desire to brief the display, you may "copy" the display to the Windows 95/NT clipboard and "paste" it directly into a presentation package such as Microsoft PowerPoint. To obtain loss vs. range and or height displays (Figures 2 and 3), you simply click the right mouse button on the coverage display.

Figure 4 shows the coverage for an airborne transmitter in the presence of an elevated duct; Figure 5 shows the simultaneous surface-based radar coverage and ESM vulnerability; and Figure 6 shows the UHF communication assessment. Note also the three earth surface depictions: dual curved, curved, and flat.

In addition to coverage displays, the effects of radar cross section variability as a function of viewing angle, ship displacement, ship height, and range are combined with the APM capabilities of range-dependent environments and terrain to produce a bar graph display (Figure 7) of detection for five classes of ship targets. These classes range from small (a patrol boat) to a very large warship (aircraft carrier). The viewing angle variability is displayed as subbars within each ship class. These angles are labeled minimum, maximum, and average, corresponding to bow, beam, and quarter.

EM Systems Database

AREPS is an unclassified program and, as such, does not include a pre-established EM system parameter database. Users are solely responsible for creating a system parameter database appropriate to their situation. To assist in this task, a database creation and maintenance capability is provided that uses fill-in-the-blank forms. Figure 8 shows such a form

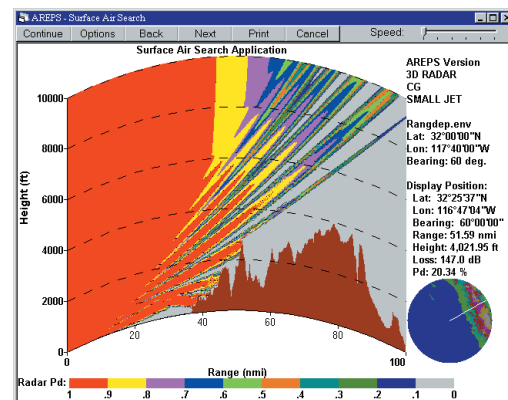


FIGURE 1. AREPS radar probability of detection coverage display.

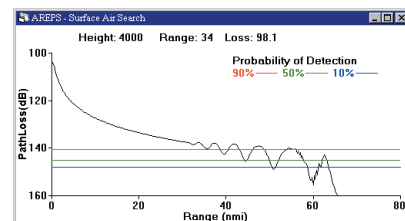


FIGURE 2. AREPS loss vs. range display.

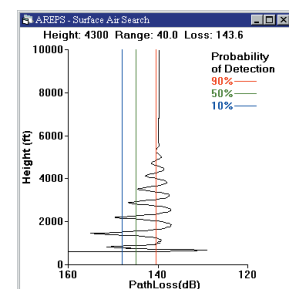


FIGURE 3. AREPS loss vs. height display.

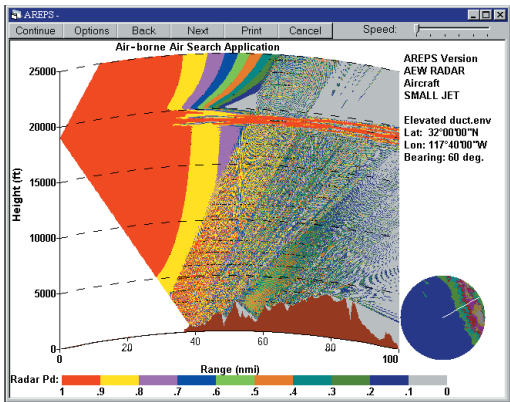


FIGURE 4. AREPS airborne air-search application.

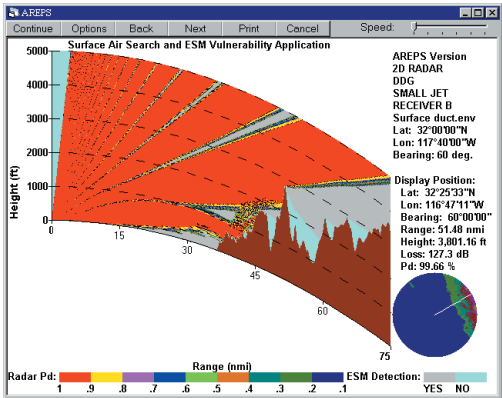


FIGURE 5. AREPS radar probability of detection and ESM vulnerability application.

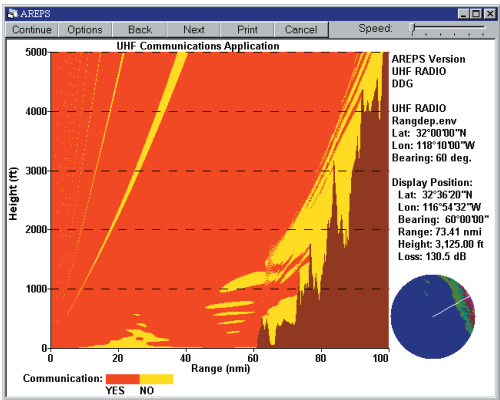


FIGURE 6. AREPS communications application.

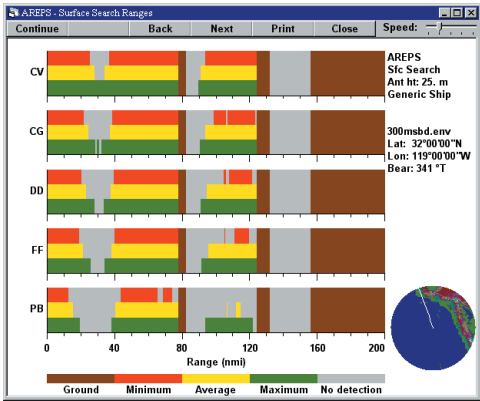


FIGURE 7. Surface-search range tables.

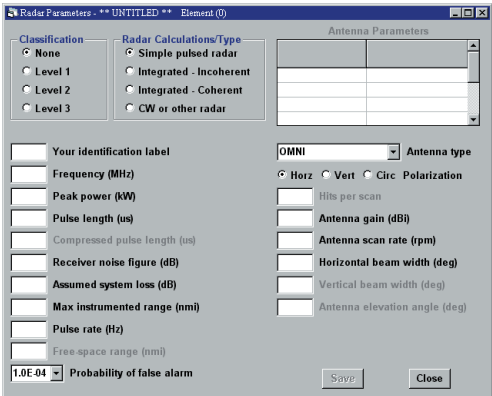


FIGURE 8. AREPS radar system input window.

for a radar system. As one navigates the form, input prompts, parameter limits, and other guidance are displayed in a status bar located at the bottom of the window.

AREPS capabilities include antenna radiation patterns of specific system height-finder antennas and a user-defined antenna pattern. Detection threshold calculations include radars using incoherent and coherent integration techniques.

In addition to pulsed radar systems, users may enter continuous wave and other non-pulsed systems, UHF and VHF communications systems, ESM receivers, and radar target descriptions.

Terrain Data

AREPS derives its terrain height data primarily from the Digital Terrain Elevation Data (DTED) provided by the National Imagery and Mapping Agency (NIMA), available either on CD-ROM or from the NIMA Internet homepage. DTED data are provided in level 0, level 1, and level 2 formats. Level 0 data spacing is 30 arc seconds in horizontal resolution (approximately 1 km). DTED level 0 data are unlimited distribution and may be obtained directly from NIMA's Internet homepage. DTED level 1 data spacing is 3 arc seconds in horizontal resolution (approximately 100 m). Level 2 data spacing is 1 arc second in horizontal resolution (approximately 30 m). Level 1 and 2 data are limited distribution. DTED data are not and may not be distributed with AREPS. For ease of input when using DTED CD-ROMs, users need only specify the latitude and longitude location of their transmitter. The AREPS program will determine which CD-ROM is required, prompt to insert the CD-ROM into the drive, and automatically extract the terrain data needed.

In addition to terrain elevations, the APM allows for the specification of range-dependent surface conditions should users be concerned about surface types for vertically polarized antennas. AREPS uses the surface conditions as defined by the International Telecommunication Union, International Radio Consultative Committee (CCIR). These conditions are provided by plain-language descriptors, selected from a drop-down menu.

Environmental Input

Atmospheric data may be derived from World Meteorological Organization (WMO) upper air observations. The entry of environmental data into AREPS has been completely automated by using the capabilities of the Windows 95/NT operating system. Within normal naval message traffic, WMO-coded radiosonde messages are routinely available. Figure 9 shows such a message.

Users need only locate the message (for a ship, the message is usually available on the ship's local area network); open the message file using any ASCII text

```
FM COMSIXTHFLT
TO OCEANO EAST
USS GEORGE WASHINGTON
USS ARTHUR W RADFORD
USS CONOLLY
USS GUAM
BT
SUBJ/UPPER AIR OBSERVATION //
RMKS/ 1. UUA 77003 99424 10053 18025 99018 17822 29023 00171 18258 31535 92838
16461 32022 85554 13464 31029 70169 05272 31032 50581 13764 29033 40747 25976 30041
30949 421// 30548 25069 // 88999 77999

UUBB 77005 99424 10053 18025 00018 17822 11989 19063 22845 13466 33835 14268
44817 13069 55// // 66771 10467 77754 09667 88// // 99731 08874 11730 08873 22// //
33707 06073 44578 07359 55551 09757 66540 10158 77539 09558 88511 12369 99463 18546
11429 22563 22414 24760 33406 25373 44381 27780 55258 505// 41414 12345 21212 00018
29023 11012 31532 22002 31535 33934 32022 44826 31030 55// // 66718 30528 77496
29034 88258 31049
BT
NNNN
```

FIGURE 9. WMO radiosonde message from Commander, Sixth Fleet.

editor (e.g., Notepad) provided with Windows 95/NT; "copy" the text to the Windows clipboard; and "paste" it into the Import WMO Code window of AREPS (Figure 10).

All extraneous text is filtered; the message is decoded; and a height vs. M-unit profile is automatically created. Should the observation be from a sea-based platform, the surface temperature and humidity are used to calculate a neutral-profile evaporation duct profile, and this profile is appended to the upper air portion of the observation. If surface observations are available, users may override the neutral profile and include full stability dependency.

It is not always necessary to have access to a local area network for the WMO observation. Many shore organizations and ships post their local radiosonde observations on their Internet or SIPRNET (Secure Internet Protocol Router Network) homepage. Once such a homepage is found for the user's particular area of interest, the WMO report may be copied to the Windows 95/NT clipboard directly from the browser (such as Netscape or Microsoft Internet Explorer), and then pasted into the Import WMO Code window. For military users, WMO reports are also available from the Fleet Numerical Meteorology and Oceanography Center (METOC) by using the Joint METOC Viewer (JMV) and/or the METOC Broadcast (METCAST) client.

For those without access to observational data in the WMO format, AREPS contains options to import observational data in a generic column format. Should real-time data be unavailable, AREPS contains a climatology of ducting conditions taken from 921 observing stations worldwide.

With the release of AREPS version 3.0, environmental data may now be obtained from mesoscale numerical meteorological models such as the Coupled Ocean and Atmosphere Mesoscale Prediction System (COAMPS). Thus, for the first time, predictions of systems' performance based on future atmospheric conditions are possible, giving the operator or the tactical decision-maker a valuable tool for mission planning.

Distribution and Support

AREPS is configured for Defense Information Infrastructure Common Operating Environment (DII COE) compliance and has been submitted as a Global Command and Control System-Maritime (GCCS-M) segment. We also provide distribution and technical support for the AREPS program. Distribution is provided on CD-ROM through U.S. mail or by direct download of the program from our Internet homepage (<http://sunspot.spawar.navy.mil>). In addition to the program software, our homepage includes help topics, frequently asked questions, and program service packages.

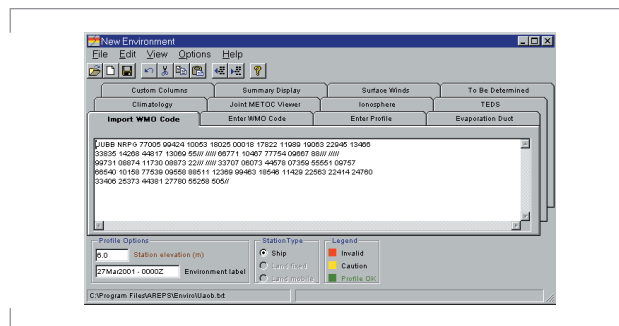


FIGURE 10. Import WMO code tab for new environmental input.



Wayne L. Patterson

MS in Meteorology, Naval Postgraduate School, 1977

Current Research: Implementation of advanced atmospheric electromagnetic wave propagation models into fleet operational systems; mesoscale meteorological model interfaces to propagation models.

A Passive Ranging Technique for Objects within the Marine Surface Layer

Stephen Doss-Hammel

SSC San Diego

INTRODUCTION

Infrared Search and Track (IRST) systems are designed to operate within the marine atmospheric surface layer. This environment can be difficult for radar systems. A reliable passive infrared (IR) system has the potential to provide useful target detection data.

However, the near sea surface environment can also distort images in the infrared. In particular, refraction effects have a strong effect on IR systems, and the occurrence of mirages is not uncommon. This report describes work to exploit one type of mirage, the inferior mirage, to determine range and height of the source creating the mirage image.

REFRACTIVE EFFECTS AND RAY-TRACE TECHNIQUES

The primary computational tool chosen for the analysis of refractive effects was a widget-based simulator called IRWarp that predicts refractive effects [1]. IRWarp uses meteorological conditions as input data for a ray-trace module [2]. The ray-trace data are used to generate detailed information about geometrical transformations induced by the propagation environment.

The ray-tracing method used within IRWarp is from a model by Lehn [3]. The radius of curvature r of a ray is given by:

$$r = \frac{nT^2}{\alpha(\lambda)(T\rho g + p dT/dz)} \quad (1)$$

where T = absolute temperature, ρ = density, p = pressure, g = gravitational acceleration, n = refractive index, and $\alpha(\lambda) = (77.6 + 0.584/\lambda^2) \times 10^{-6}$ for wavelength $= \lambda$. It is also assumed that the ray slope does not exceed 10 milliradians.

The formulation in Eq. (1) applies to visible and infrared wavelengths. Pressure is relatively constant for the measurements made, and the prime determinant of the radius of curvature of near-horizontal rays was the vertical temperature gradient. The ray-trace algorithm first defines the vertical temperature profile as a set of discrete layers, each with a characteristic temperature gradient and refractivity gradient. A characteristic radius of curvature is then assigned to each layer using Eq. (1).

The vertical temperature profile is based upon a surface-layer similarity theory developed by Monin and Obukhov. For the current study, an approach was followed based upon bulk methods for calculating turbulence

ABSTRACT

Infrared Search and Track (IRST) systems are important to the surface Navy for the detection of low-flying missile threats. Infrared signals propagating within the marine atmospheric surface layer are frequently distorted by strong vertical fluxes. One particular distortion that occurs commonly is the sub-refractive mirage. During sub-refractive mirage conditions, an imaging sensor or camera will record two distinct images of a single point source. A sub-refractive mirage image can be exploited to provide both height and range information. A technique for passive ranging is described, and a case study using field test data is presented as an example of the concept.

parameters described by Davidson et al [4]. Field measurements were taken at the sea surface, and at a reference height, and these values were used to determine the particular values of the scaling parameters. Thus, the sea surface temperature would be T_0 , and the temperature $T(z)$ at a height z above the water surface would be given by

$$T = T_0 + T_* \left[\frac{\ln(z/Z_{0T} - \psi_T(z/L))}{\alpha_T k} \right]$$

where Z_{0T} is the roughness length for the temperature profile, T_* is the potential temperature scaling parameter, and α_T is the ratio of heat transfer to momentum transfer at the surface. L is the Monin–Obukhov length, and $\psi_T(z/L)$ is a stability correction function.

A ray trace can be generated from the temperature profile by determining a characteristic radius of curvature for each horizontal layer using Eq. (1). Figure 1 displays the traced rays from the ray-trace algorithm for a coordinate system transformed so that the sea surface is the flat x-axis. The figure shows a ray-trace generated from field test temperature profile data. The air–sea temperature difference was ≈ -3.5 K. The number of rays has been reduced to make the graphic more legible. The apparent kinks in some of the more sharply bent rays are an artifact; the actual path for the ray is a carefully determined smooth curve, but points on the path are saved only intermittently as needed for the calculation.

An atmospheric surface layer for which the air–sea temperature difference is negative exhibits a crucial feature: the rays form a local coordinate system starting at some point downrange. The logarithmic temperature profile ensures that lower elevation rays are deflected to intersect upper elevation rays. The existence of a locally non-degenerate coordinate system implies that in some region of range–height space there exists a one-to-one correspondence with an upper elevation–lower elevation pair that is unique to that point.

TRANSFORMING IMAGE ELEVATION TO HEIGHT-RANGE DATA

The set of rays tracing the propagation path defines an envelope. The ray envelope has an intersection structure with a set of constant-height surfaces (see Figure 1) at heights of 4, 6, 8, 10, 12, and 14 m. A ray traced from the receiver intersects a given constant-height surface either once, twice, or not at all. The intersection structure of the constant-height surfaces with the ray-trace envelope induces a transformation.

To understand transformation more completely, consider Figure 2. The term "isomet" (isomet surface \equiv surface of constant height) is used to

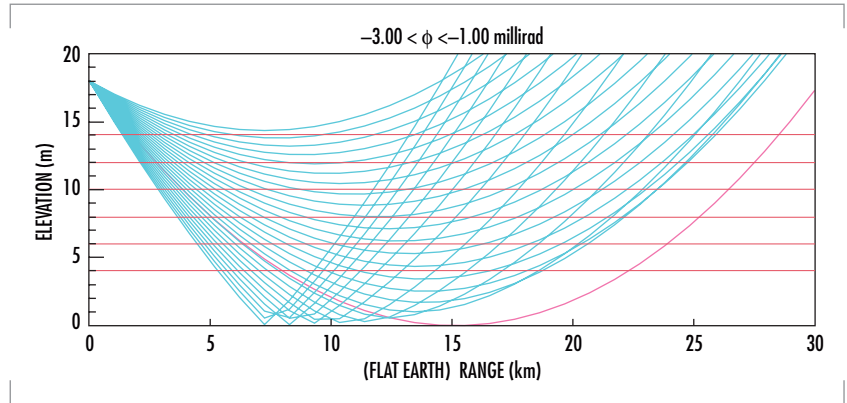


FIGURE 1. The vertical scale is in meters above the surface, while the horizontal scale is in kilometers downrange from the sensor (height = 18 m). The six red horizontal lines indicate level surfaces. The rays that appear to be reflecting from the x-axis are actually refracted.

refer to the contour curves representing the intersection set between a constant-height surface and the ray-trace envelope shown in Figure 2. Each of the isomets in Figure 2 displays a similar form. The vertical axis shows angular displacement from the horizontal tangent plane at the sensor. The horizontal axis shows range.

The graph of a single isomet can be interpreted by imagining a source confined to one of the isomet surfaces (for example, the 14-m isomet) and moving toward the sensor from the 30-km range. At ≈ 26 km, the source appears over the horizon as a single point that immediately splits into two images. As seen through an imaging sensor, for example, one image decreases in angular elevation, and the upper image increases in angular elevation as the source moves closer in range. At ≈ 13 km, the lower image descends below -3 milliradians; in terms of the imaginary sensor, it has descended beneath the lowest edge of the sensor focal plane. The (now solitary) upper image continues to rise to the upper edge of the sensor field of view. Within the last 6 km, the source is seen to rapidly move from near the top edge to disappear below the bottom edge.

This form for the 14-m isomet is characteristic of all the isomet contours for surfaces of height less than the sensor height. When the isomet surface height is greater than sensor height, an inbound upper image disappears across the upper boundary, and never re-crosses from top to bottom.

The key to a deduction of height and range from angular elevation information is the utilization of those portions of an isomet for which two values of elevation correspond to a single range value. Thus, for the 14-m isomet, ranges between 13 km and 25 km correspond to two distinct elevation values. This indicates that it is possible to find a one-to-one correspondence between a pair of elevation angles, and a height-range pair.

Thus, the central result in this paper is the transformation shown in Figure 3. When a sensor detects two images, the elevations of the lower and upper images can be plotted as a point in Figure 3, and the height and range of that point can be read from the inner coordinate system. To say it differently, the figure contains the transformation that takes two elevation measurements as input, and generates as output both height and range of the source or target. In terms of coordinate systems, the rectilinear lower elevation vs. upper elevation coordinate system is transformed to the distorted, curvilinear height vs. range coordinate system.

Consider as an example an imaging sensor system with a telescope that detects a source in a

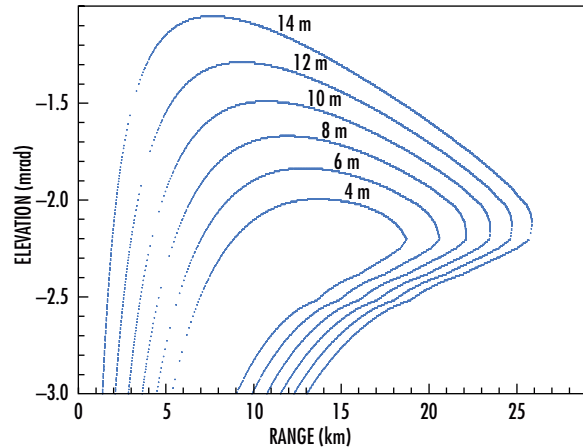


FIGURE 2. A series of isomets at the heights of 4, 6, 8, 10, 12, and 14 m. For a given range value, each isomet defines either 0, 1, or 2 corresponding elevation values.

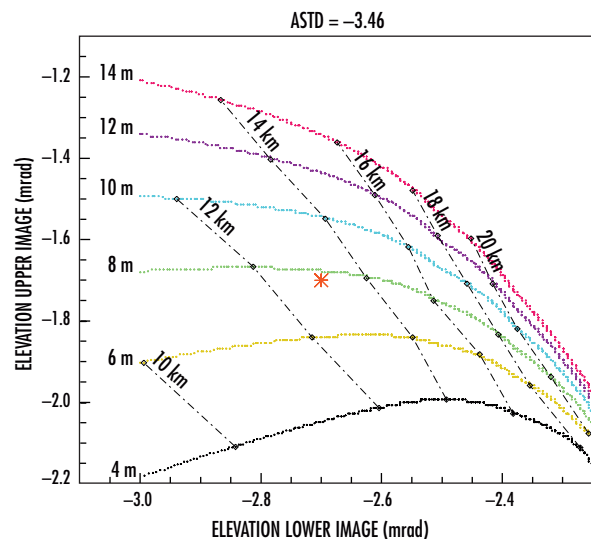


FIGURE 3. The transformation that is implied by the data in Figure 2. The same information is shown here, but restricted to the portions of the isomets that are dual-valued. The point $(\theta_{\text{lower}}, \theta_{\text{upper}}) = (-2.7, -1.7)$ is plotted as an example, and it transforms to range ≈ 13 km, height ≈ 7.5 m.

sub-refractive mirage regime. The two elevations can be determined from the imaging frame: suppose $(\theta_{lower}, \theta_{upper}) = (-2.7, -1.7)$. This example is plotted in Figure 3. Using the transformation, the actual range and height can be read out from the transformed coordinate system, yielding range ≈ 13 km and height ≈ 7.5 m.

SUMMARY

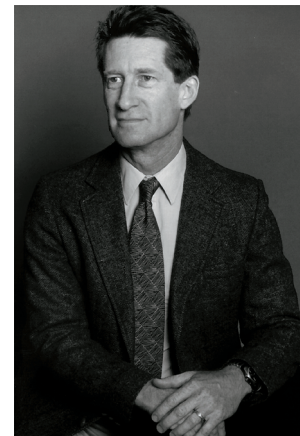
Sub-refractive conditions are quite common for the marine atmospheric surface layer. These conditions cause mirages that appear at two different elevations. These two elevations can be transformed by means of a ray-trace technique to yield height and range information.

The usable range for the particular example presented here is from 10 or 12 km out to ≈ 20 km. Note that the range limits for effective range-finding are determined by the intensity of the sub-refractive conditions. As air-sea temperature difference $T_{air} - T_{sea}$ becomes more negative, the range domain for which two images occur increases in extent by moving the point of first appearance of two images closer to the sensor. Conversely, as air-sea temperature difference $T_{air} - T_{sea}$ becomes less negative and closer to zero, the range domain for which two images occur decreases in extent; the first appearance of two images occurs at a point farther away from the sensor.

Numerous issues remain to be explored. It is necessary to define the limits of applicability for the method. It is also necessary to establish a mathematical foundation for assumptions made concerning the behavior of surfaces and the intersections between them. Furthermore, the method uses an implicit assumption of homogeneity: the full propagation range is characterized by one vertical profile. This appears to be a reasonable assumption for the sub-refractive case, but this also must be carefully examined. To make practical use of the passive ranging technique, it is important to calculate the limits to the angular elevation resolution.

REFERENCES

1. Hammel, S. and N. Platt. 1995. "Topological Description of Mirage Effects," *Proceedings of the Vision Geometry Conference*, SPIE, vol. 2573, pp. 398–406.
2. Platt, N., S. Hammel, J. Trahan, H. Rivera. 1996. "Mirages in the Marine Boundary Layer—Comparison of Experiment with Model," *Proceedings: IRIS Passive Sensors*, vol. 2, pp. 195–210.
3. Lehn, W. 1985. "A Simple Parabolic Model for the Optics of the Atmospheric Surface Layer," *Applied Math. Model*, vol. 9, p. 447.
4. Davidson, K., G. Schacher, C. Fairall, and A. Goroch. 1981. "Verification of the Bulk Method for Calculating Overwater Optical Turbulence," *Applied Optics*, vol. 20, p. 2919.



Stephen Doss-Hammel

Ph.D. in Applied Mathematics,
University of Arizona, 1986
Current Research: Atmospheric
optics and dynamical systems.

Silicon-on-Sapphire Technology: A Competitive Alternative for RF Systems

Isaac Lagnado and Paul R. de la Houssaye
SSC San Diego

S. J. Koester, R. Hammond, J. O. Chu, J. A. Ott,
P. M. Mooney, L. Perraud, and K. A. Jenkins
IBM Research Division, T. J. Watson Research Center

INTRODUCTION

Device-quality, thin-film silicon-on-sapphire (TFSOS), obtained by Solid Phase Epitaxy (SPE), has achieved truly outstanding results that are incorporated into present and future high-performance products, such as phase-locked loop integrated circuits (ICs) for wireless communications [1], single-chip Global Positioning System (GPS) receivers, and analog to digital converters (A/DCs) for space applications.

EXPERIMENTAL STUDY

Table 1 shows the measured performance of a front-end receiver at 2.4 GHz; and voltage-controlled oscillators (VCOs), frequency dividers, and tuned amplifiers at >20 GHz. The n-metal oxide semiconductor (MOS) VCO at 26 GHz [2] has the highest tuned frequency ever achieved among complementary metal-oxide semiconductor (CMOS) VCOs.

Concomitantly, recent advances in SiGe epitaxial growth technology indicate that SiGe-based strained-layer modulation-doped field-effect transistors (MODFETs) may be promising alternatives to III-V metal-semiconductor field-effect transistors (MESFETs) and high-electron mobility transistors (HEMTs) for future high-speed analog communications applications. Electron and hole mobilities well in excess of bulk Si mobilities can be realized in tensile-strained Si quantum wells (QWs) [3] and compressive-strained SiGe [4] or pure Ge QWs [5], respectively. Note that p-MODFETs have demonstrated dc and RF performance figures comparable to n-MODFETs, suggesting the possibility of very-high-speed complementary operation [6], a capability not available in current III-V technology. In this work, we have applied this knowledge to the growth of SiGe relaxed buffer layers and the fabrication of SiGe strained-layer MODFETs on TFSOS. One of the benefits of using insulating substrates, such as sapphire, is the potential solution to the reduction of high losses seen in the microwave frequency regime due to the conducting nature of the silicon substrate. Here, we demonstrate the development of the epitaxial growth of high-mobility, modulation-doped, composite-channel heterostructures on silicon-on-sapphire (SOS) substrates, and describe the resultant outstanding RF characteristics of ≤ 100 nm T-gate p-MODFETs fabricated on these layer structures.

The composite-channel heterostructure device has the basic structure shown in Figure 1. This layer structure was grown on an SOS substrate

ABSTRACT

We investigated the formation of high-performance, device-quality, thin-film silicon (30 to 50 nm) on sapphire (TFSOS) for application to millimeter-wave communication and sensors. The resulting TFSOS, obtained by Solid Phase Epitaxy (SPE), and the growth of strained silicon-germanium (SiGe) layers on these TFSOS demonstrated enhanced devices and, hence, integrated-circuit performance not achieved previously. We fabricated 250-nm and 100-nm T-gated devices with noise figures as low as 0.9 dB at 2 GHz and 2.5 dB at 20 GHz, with G_a of 21 dB and 7.5 dB, respectively. 250-nm devices resulted in distributed wideband amplifiers (10-GHz bandwidth [BW], world record) and tuned amplifiers (15-dB, 4-GHz BW). 100-nm devices produced voltage controlled oscillators (VCOs) (25.9-GHz), 30-GHz frequency dividers. We obtained f_t (f_{max}) of 105 GHz (50 GHz) for n-channel and 49 GHz (116 GHz, world record) for p-MODFETs (strained $Si_{0.2}Ge_{0.8}$ on a relaxed $Si_{0.7}Ge_{0.3}$ hetero-structure). This paper details our investigation and provides cost comparisons with competing technologies.

TABLE 1. Measured performance of various components of a front-end receiver designed to operate near 2.4 and 18 GHz.

Measured	Operating Frequency	Gain	NF (50 Ω)	IP3 (output)	Power@Vdd
LNA	2.4 GHz	11 dB	2.2 dB	14 dBm	13.2 mW@1.5V
LNA (HEMT, 2 stages)	1.4–2.6 GHz	25 dB	2.3 dB	15 dBm	
Mixer	Center = 2.4 GHz IF = 250 MHz	-5 dB		5 dBm	8.4 mW @1.5V
Mixer (HBT)	1.4–2.6 GHz	-4 dB	15 dB	0 dBm	
VCO	25.9 GHz 0.6-GHz tuning range		-106 dBc/Hz (phase noise)		24 mW @1.5V
Frequency Divider	1.5–20 GHz 5.9–26.5 GHz				29.5 mW (Core) <20 mW (Core)
Tuned Amplifier	23 GHz 4-GHz Bandwidth	6–7 dB			

as well as a bulk Si control wafer. The devices had a gate length, L_g , of 100 nm.

The room-temperature Hall mobility and sheet carrier density [7] of the composite-channel layer structure grown on an SOS wafer were 800 to 1200 cm²/Vs and 3.1–2.5 $\times 10^{12}$ cm⁻² at room temperature, respectively, as shown in Figure 2.

The room-temperature output (transfer) characteristic for 100-nm gate-length devices showed practically no difference between devices fabricated on SOS and Si control devices (Figure 3).

Figure 3 also shows that the best SOS transistor had a maximum extrinsic transconductance of 377 mS/mm, which is, to the authors' knowledge, the highest ever reported for alloy-channel p-MODFETs. The device had a corresponding output conductance of 25 mS/mm leading to a maximum dc voltage gain of 15. The only apparent degradation of the device performance caused by the SOS substrate was roughly an order of magnitude higher gate-leakage current compared to the Si monitor, a result that we again attribute to the increased defect density of the SOS wafers.

Figure 4 shows frequency-dependent plots [8] of the forward current gain ($|b_{21}|^2$) and the maximum unilateral gain (MUG) for a 0.1 \times 50 μm^2 p-MODFET on SOS. Values of $f_T = 49$ GHz and $f_{\text{max}} = 116$ GHz were obtained after de-embedding the contact pads; the latter value being the highest f_{max} ever reported for a SiGe p-MODFET. Figures 5 and 6 illustrate the fact that f_t and f_{max} saturate at a very low bias voltage; f_{max} reached 100 GHz at $V_{\text{ds}} \sim 0.6$ V and remained sustained over a wide bias range. In Figure 7, the small-signal parametric model reveals non-negligible capacitances (C_{pg} and C_{pd}) caused by the incomplete removal of the SiGe buffer layer and Si film in the isolation regions. Through the agreement of the raw data and extracted values, Figure 8 demonstrates the accuracy

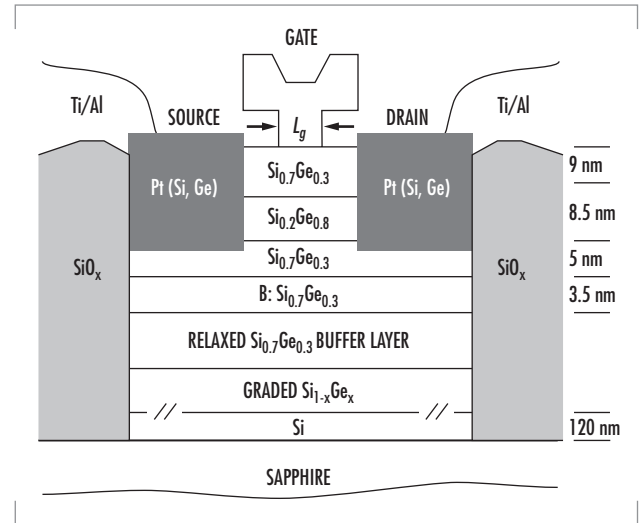


FIGURE 1. Cross-sectional diagram of epitaxial layer structure and p-MODFET device design.

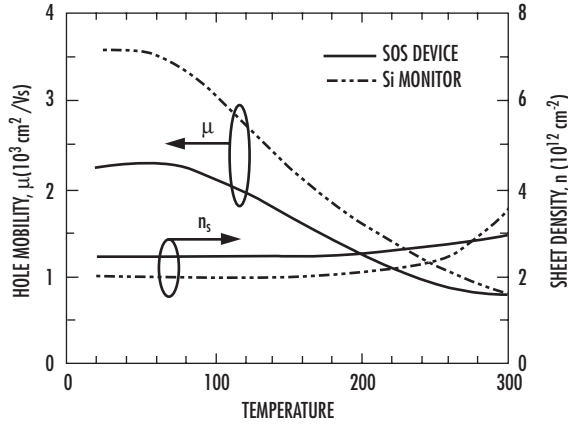


FIGURE 2. Hole mobility and sheet density vs. temperature for composite-channel layer structures grown on SOS and Si control wafers.

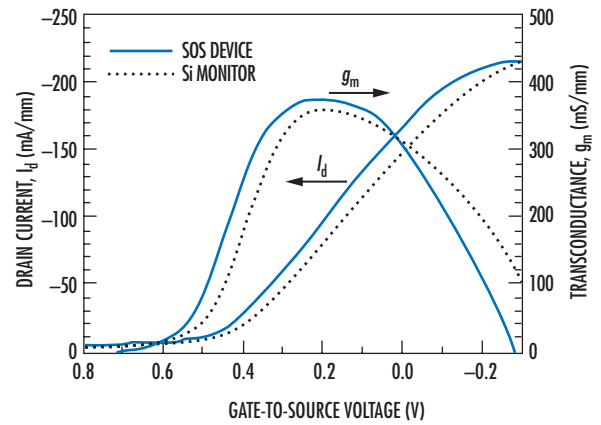


FIGURE 3. Comparison of 0.1- μm composite-channel p-MODFETs on Si and SOS. The bias voltage is $V_{ds} = -0.6$ V.

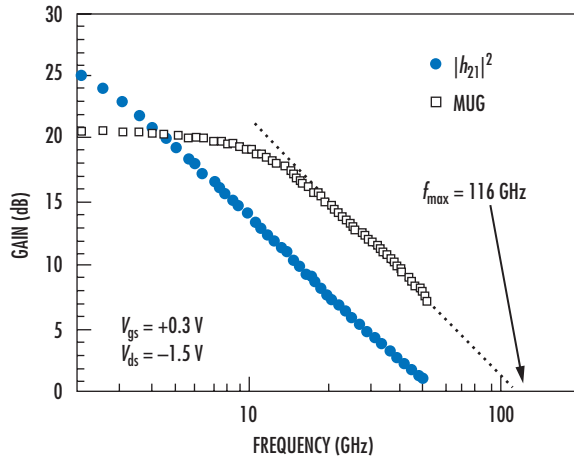


FIGURE 4. Plot of $|h_{21}|^2$ and MUG vs. frequency for a $0.1 \times 50 \mu\text{m}^2$ composite-channel p-MODFET on SOS. Values of $f_t = 49$ GHz and $f_{\text{max}} = 116$ GHz are obtained after open-circuit de-embedding.

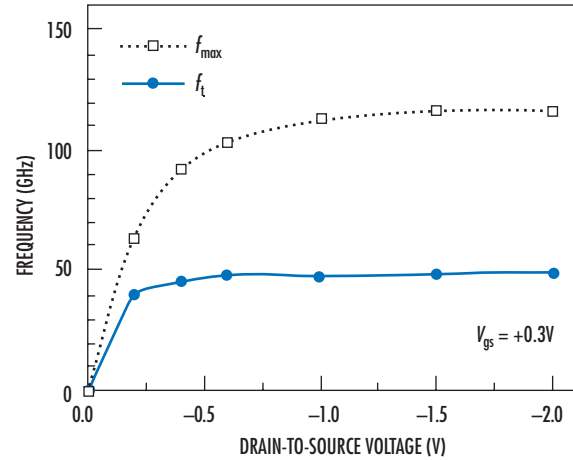


FIGURE 5. Bias dependence of f_t and f_{max} . f_t and f_{max} saturate at a very-low-bias voltage.

of the preliminary device model created from these results. Design of circuits from this model should demonstrate the unprecedented potential of SiGe/TFSOS technology.

Table 2 lists the comparative cost of competing technologies [9, 10, and 11] for the manufacturer and the user.

CONCLUSION

The incorporation of strained SiGe heterostructures on thin-film silicon-on-sapphire (the device-quality Si film obtained either through SPE or layer bonding) for n- and p-FETs, characterized by superior transport carrier properties, high dynamic performances (f_t , f_{max}), and low noise at high frequencies (Figure 9) will enable an entirely new technology. The

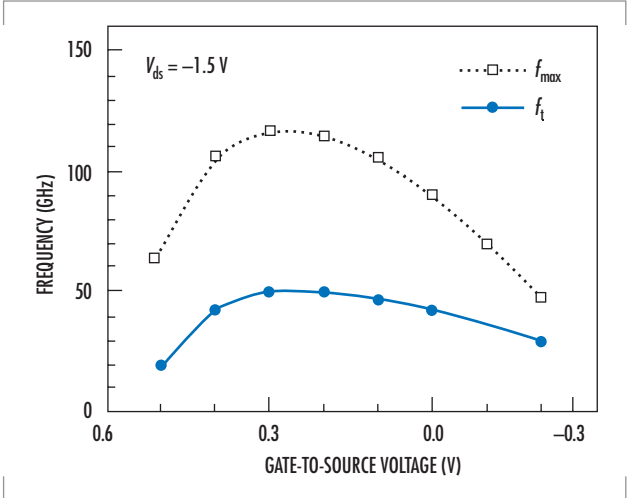


FIGURE 6. Bias dependence of f_t and f_{max} . High f_{max} sustained over wide bias range.

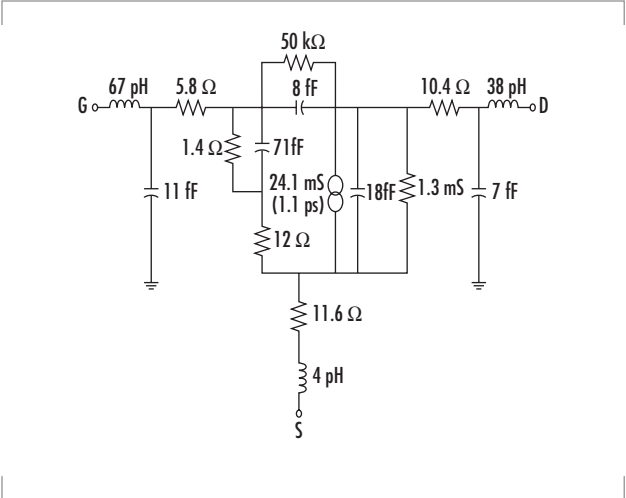


FIGURE 7. Small-signal equivalent circuit. Small-signal parameters reveal non-negligible capacitances (C_{pg} and C_{pd}) from unremoved SiGe buffer layer in isolation regions.

TABLE 2. Relative cost of different technologies as seen by (A) the manufacturer and (B) the user. The difference is due to the different profit margins available to the companies as determined by what the market will bear.

A. Manufacturing Cost [9, 10]				
Si Bulk CMOS	SOS CMOS	Bipolar Si & SOS SiGe	GaAs MESFET	HBT(GaAs)
1	1.3	3.5	3.5–7	10

B. User Cost [11]	
Technology	Cost per sq. mm (\$US)
Silicon CMOS	0.01
SiGe epitaxy	0.60
GaAs epitaxy	2.00
InP epitaxy	10.00
Tokyo real estate	0.01

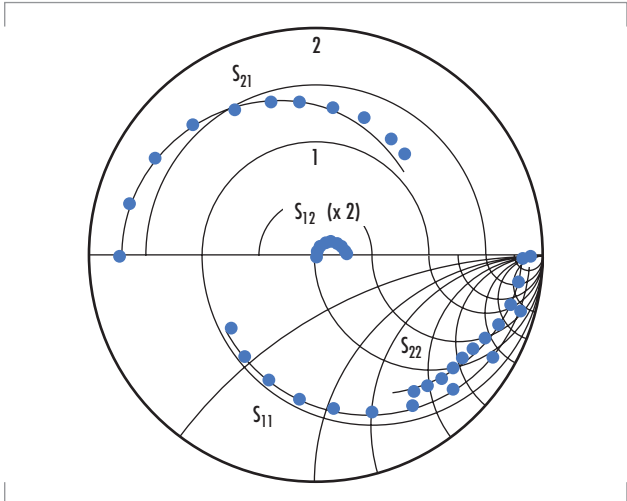


FIGURE 8. Comparison of s-parameter. Good agreement between raw data (points) and extracted values (lines).

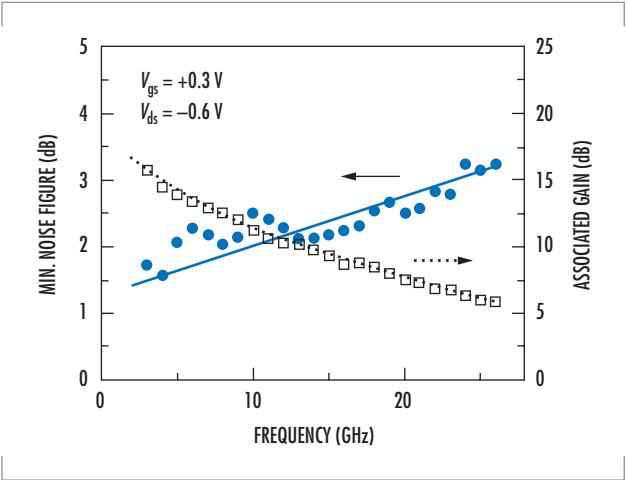


FIGURE 9. Noise parameter results. Values of $F_{min} = 2.5$ dB and $G_a = 7.5$ dB obtained at 20 GHz. Low-frequency noise dominated by gate-leakage current.

resulting impact of the combined TFSOS and SiGe technology on the marketplace, both nationally and internationally, will be quite revolutionary because no other material can provide a complementary technology as efficiently, from either the technical or economic aspect.

REFERENCES

1. Peregrine Semiconductor Corporation data sheet.
2. Wetzel, M., L. Shi, K. Jenkins, P. R. de la Houssaye, Y. Taur, P. M. Asbeck, I. Lagnado. 2000. "A 26.5-GHz Silicon MOSFET 2:1 Dynamic Frequency Divider," *IEEE Microwave and Guided Wave Letters*, vol. 10, no. 10, pp. 421–423.
3. Nelson, S. F., K. Ismail, J. O. Chu, and B. S. Meyerson. 1993. "Room-Temperature Electron Mobility in Strained Si/SiGe Heterostructures," *Applied Physics Letters*, vol. 63, pp. 367–369.
4. Ismail, K., J. O. Chu, and B. S. Meyerson. 1994. "High Hole Mobility in SiGe Alloys for Device Applications," *Applied Physics Letters*, vol. 64, pp. 3124–3126.
5. Höck, G., M. Glück, T. Hackbarth, H.-J. Herzog, and E. Kohn. 1998. "Carrier Mobilities in Modulation Doped $\text{Si}_{1-x}\text{Ge}_x$ Heterostructures with Respect to FET Applications," *Thin Solid Films*, vol. 336, pp. 141–144.
6. Armstrong, M. A., D. A. Antoniadis, A. Sadek, K. Ismail, and F. Stem. 1995. "Design of Si/SiGe Heterojunction Complementary Metal-Oxide-Semiconductor Transistors," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 761–764.
7. Koester, S. J., R. Hammond, J. O. Chu, J. A. Ott, P. M. Mooney, L. Perraud, and K. A. Jenkins, I. Lagnado, P. R. de la Houssaye. 1999. "High-Performance SiGe pMODFETs Grown by UHV-CVD," *Proceedings of the 7th International Symposium on Electron Devices for Microwave and Optoelectronic Applications (EDMO 99)*, 22 to 23 November, King's College, London, England.
8. Koester, S. J., R. Hammond, J. O. Chu, P. M. Mooney, J. A. Ott, C. S. Webster, I. Lagnado, and P. R. de la Houssaye. 2000. "Low-Noise SiGe pMODFETs on Sapphire with 116-GHz f_{max} ," *58th IEEE Device Research Conference*, 19 to 21 June, Denver, CO.
9. Young, J. P. and R. Collins. 1997. Course given at *IEEE Microwave Theory and Techniques Symposium (MTT-S)*.
10. Wienau, D. 2000. "SiGe RF Technology for Mobile Communications—Technical and Commercial Aspects," *Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems: Digest of Papers*, pp. 79–82.
11. Paul, J. D. 1999. "Silicon-Germanium Strained-Layer Materials in Microelectronics," *Advanced Materials*, vol. 11, no. 3, pp. 191–204.



Isaac Lagnado

Ph.D. in Solid-State Electronics, MIT, 1970

Current Research: Thin-film silicon-on-sapphire (TFSOS); silicon germanium (SiGe) on TFSOS; application to RF mixed-signal systems on a single chip (SOC); high performance analog-to-digital converters; high-efficiency, high-power density power supplies.

Paul R. de la Houssaye

Ph.D. in Applied Physics, Stanford University, 1988

Current Research: Thin-film silicon/SiGe on sapphire device and circuit fabrication; novel microelectromechanical systems (MEMS).